Essays on Cross-Commodity Modeling in Energy Markets

Christensen, Troels Sønderby

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

ESSAYS ON CROSS-COMMODITY MODELING IN ENERGY MARKETS

BY
TROELS SØNDERBY CHRISTENSEN

DISSERTATION SUBMITTED 2019

AALBORG UNIVERSITY
DENMARK

Essays on Cross-Commodity Modeling in Energy Markets

PhD Dissertation

Troels Sønderby Christensen

Dissertation submitted December 20, 2019

Dissertation submitted:	December 20, 2019
PhD supervisors:	Assoc. Prof. Esben Høg Aalborg University
	Jesper Jung, PhD Centrica Energy Trading
PhD committee:	Professor Rasmus Waagepetersen (chairman) Aalborg University
	Professor Almut Veraart Imperial College, London
	Professor Asger Lunde Aarhus University
PhD Series:	Faculty of Engineering and Science, Aalborg University
Department:	Department of Mathematical Sciences

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-569-7

Published by:
Aalborg University Press
Langagervej 2
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2020

Til Kalle og den lille Mubbi

Preface

This thesis is the result of my enrolment in the Industrial PhD program funded by the Innovation Fund of Denmark and Centrica Energy Trading for the period from January 2017 to December 2019. The thesis consists of a brief introduction followed by five research papers. Although the material presented in each paper is self-contained, two common themes are present: 1) renewables and 2) interconnectedness, both in relation to European electricity markets. The first three papers are concerned with the effects of the increased volumetric risk for electricity market participants introduced by the increased wind power production. In the fourth and fifth paper we consider the effects of the increased interconnectedness between different market areas.

I would like to express my gratitude towards Centrica Energy Trading for giving me the opportunity to pursue the PhD degree. Especially I would like to thank my company supervisor Jesper Jung for the great support and guidance during the three-year period. The Quantitative Analytics department at Centrica Energy Trading is also thanked for the friendly atmosphere and the great discussions we have had.

A huge thanks goes to my university supervisor Esben Høg for providing excellent guidance, which already began during my master's studies. My fellow PhD students at the Department of Mathematical Sciences at Aalborg University is also thanked for contributing to a stimulating and enjoyable research environment. In this regard, I would like to thank former PhD student Anca Pircalabu for our great collaboration and our two joint papers.

I got the opportunity to visit Professor Fred Espen Benth at the University of Oslo for three months from January 2018 to March 2018. I would like to thank Fred for his inspiring way of providing guidance. I would also like to thank the rest of the members of the stochastic analysis group at the Department of Mathematics for their hospitality and for introducing me to cross-country skiing.

Troels Sønderby Christensen
Aalborg, December 20, 2019

Summary

This thesis consists of an introduction and five research papers on cross-commodity modeling in energy markets. Since all papers have a connection to European electricity markets, special attention is paid to these markets in the introductory part of the thesis.

In Papers A, B, and C we consider the mitigation of volumetric risk introduced by the increased wind power production in Europe. In Paper A we propose a seasonal copula mixture for the joint modeling of the clean spark spread and the underlying of the exchange-traded wind power futures. Assuming that the clean spark spread represents the profitability of a gas-fired power plant, the proposed model facilitates through Monte Carlo simulations the assessment of wind power futures as hedging instrument for conventional power producers, here represented by a gas-fired power plant.

The point of view is shifted in Paper B, where we take the view of 31 different wind power producers located in Germany. We propose ARMA-GARCH copula models for the joint behaviour of each pair of site-specific wind index and the underlying index of the exchange-traded wind power futures. Monte Carlo simulations from the model allow us to derive hedging strategies and to quantify the differences in hedging effectiveness of wind power futures for the considered wind power producers.

In Paper C we propose two continuous-time multivariate models for the joint behaviour of wind indexes based on Lévy processes. In addition, we put forward the suggestion of tailor-made wind power futures to facilitate a perfect hedge of the volumetric risk faced by wind power producers. Using both models, we analyze the tailor-made wind power futures for the selling and buying side, and discuss the similarities and differences between the models.

Papers D and E are concerned with the implication of the price formation in European day-ahead electricity markets, where market coupling is present. In Paper D we propose a continuous-time model for the joint behaviour of day-ahead electricity prices in two coupled interconnected markets. Estimation procedures, based on filtering techniques, are proposed, and closed-form formulas for forwards and transmission rights implied by the model are pre-

sented.

In Paper E we consider the model proposed in Paper D in more detail with regard to trading transmission rights. Key features of the model are discussed in this context. In the empirical part of the paper we asses the model pricing and hedging wise in relation to transmission rights. For comparison reasons, we compare the model to the Margrabe formula throughout Paper E due to its well-known properties.

Resumé

Denne afhandling indeholder en introduktion og fem forskningsartikler omkring krydsmodellering af råvarer i energimarkeder. Vi har særligt fokus på europæiske elektricitetsmarkeder i introduktionen af denne afhandling, da alle artikler har en forbindelse til disse markeder.

I Artikler A, B og C betragter vi håndteringen af volumenrisiko, som følge af introduktionen af den øgede elektricitet produceret af vindmøller i Europa. I Artikel A foreslår vi en sæson copula mixtur for den simultane modellering af clean spark spread og det underliggende til de børshandlede wind power futures. Vi antager, at clean spark spread repræsenterer rentabiliteten af et gaskraftværk. Den foreslåede model faciliterer dermed, gennem Monte Carlo simulationer, vurderingen af den risikominimerende effekt af wind power futures for konventionelle elektricitetsproducenter repræsenteret ved et gaskraftværk.

I Artikel B skifter vi synspunkt til vindmøller lokaliseret i Tyskland. Vi foreslår ARMA-GARCH copula modeller for den simultane opførsel af hvert par af vindmølle indeks og det underliggende indeks af de børshandlede wind power futures. Ved hjælp af Monte Carlo simulationer fra modellerne udleder vi risikoafdækkende strategier og kvantificerer forskellene i risikominimeringen som følge af wind power futures for de betragtede vindmøller.

I Artikel C foreslår vi to kontinuerttids multivariat modeller for den simultane opførsel af vindindeksene baseret på Lévy processer. Yderligere foreslår vi skræddersyede wind power futures for at imødekomme en perfekt risikominimering af den volumenrisiko som vindmølleejere skal håndtere. Begge modeller bliver brugt til at analysere de skræddersyede wind power futures kontrakter fra både et køber- og sælgerperspektiv, og ligheder og forskelle på de to modeller bliver diskuteret.

Artikler D og E omhandler implikationen af prisformationen i europæiske day-ahead elektricitetsmarkeder, hvor markedskoblingen er til stede. I artikel D foreslår vi en kontinuerttids model for den simultane opførsel af day-ahead elektricitetspriser i to sammenkoblede markeder. Vi foreslår estimationsprocedurer baseret på filtreringsteknikker, og formler for terminskontrakter og transmissionsrettigheder på lukket-form.

I Artikel E betragter vi modellen foreslået i Artikel D med et mere detaljeret fokus på handlen med transmissionsrettigheder. Hovedkarakteristika af modellen er diskuteret i denne kontekst. I den empiriske del af artiklen vurderer vi modellens evne til at prisfastsætte og risikohåndtere transmissionsrettigheder. Margrabe formlen er brugt gennem hele artikel E som sammenligningsgrundlag grundet dens velkendte egenskaber.

Contents

Part I

Introduction

Background

The overall theme of this thesis is cross-commodity modeling limited to energy markets. Cross-commodity modeling refers here to the joint modeling of different commodities such as natural gas and electricity, but also the joint modeling of the same commodity in different areas such as wind power production at distinct geographical locations. This introduction serves to describe how both variations of cross-commodity modeling is touched upon in the included research papers. Since all papers are linked to European electricity markets, we introduce the papers in the context of these markets.

Markets for Electricity in Europe

The liberalization of energy markets globally has resulted in severe changes in the whole energy sector. From a European perspective, the change from monopolized to liberalized electricity markets was the consequence of the introduction of a single European market in the beginning of the 1990s. One intention of the European Union was to provide free movement of goods and services, though differences in electricity prices could potentially incentivize companies to relocate based on the electricity costs associated with each member state (Heddenhausen (2007)). The plan of a single pan-European wholesale electricity market was thus initiated.

A common target model for the member states of the European Union was finalized in 2009, including a specification of the types of electricity markets needed to foster the integration of electricity markets (see Danish Energy Agency (2018)):

- A central element in this target model is the so-called day-ahead market, where market participants can trade electricity for each hour of the following day. Buyers and sellers of day-ahead electricity have to submit their orders for all hours of the following day before 12:00 CET. For each hour, a uniform pricing method is applied resulting in one price where the supply and demand curves intersect. Buyers willing to pay at least this intersection price will get their bid fulfilled, while sellers

willing to sell electricity to at most this intersection price will get their offer fulfilled.

- To allow market participants to hedge their exposure to the day-ahead prices, forward markets are also part of the target model, with the day-ahead price being the underlying of the corresponding forward contract. Compared to traditional financial markets, a forward contract in electricity markets is characterized by a delivery period instead of a single delivery point in time. Opposite the day-ahead market, buyers and sellers are matched continuously until delivery of the forward contract.

- To allow market participants to hedge their exposure to physical congestion constraints in interconnected markets, markets for long-term transmission rights (also referred to as cross-border capacities) are also part of the target model. Different ways of introducing such instrument has been applied.

  - A physical transmission right (PTR) gives the owner the right to transfer an amount of electricity for a pre-specified period of time from one market area to another market area[1]. To use this right, the owner must daily notify whether she will use the right or not for the following day (this is also termed nomination of the transmission right). In Europe, most PTRs are sold under the "use-it-or-sell-it" (UIOSI) principle. Given a PTR entitling the right to transfer electricity from market A to market B, the UIOSI principle implies that the owner of the PTR not nominating the PTR for a given day, will be compensated with the day-ahead price in market A subtracted from the day-ahead price in market B, if the spread is positive. As with the forward contract, a transmission right is characterized by a delivery period; e.g. a day, a week, etc.
  - A financial transmission right (FTR) is in Europe typically defined as a basket of European spread options where the underlying is the day-ahead price spread between interconnected markets. Thus, the compensation of an FTR is equivalent to the compensation of the PTR under the UIOSI principle, if the owner of the PTR do not nominate it.

  A uniform pricing method is applied in both cases.

[1] Each member state in the European Union usually constitute a market area with a common electricity price for the corresponding delivery period; e.g., the day-ahead price is valid for each member state. However, there are exceptions, where the member state is divided in more than one market area, such as Denmark.

The Role of Renewable Electricity Production

Electricity produced from renewables has increased tremendously through recent years. In the European Union the total installed capacity of wind and solar increased from 12.9 GW in 2000 to 218.2 GW in 2014, which should be compared to a total installed capacity of the 28 member states of 628 GW in 2000 and 977 GW in 2014 (the Commission Expert Group (2017)).

This increase in renewables has lead to decreases in wholesale electricity prices all over Europe. According to European Commission (2016), electricity prices fell by almost 70% from 2008 to 2016. The price formation of the day-ahead market, introduced in the previous section, is illustrated in Fig. 1, where the demand is assumed inelastic. The renewable power production replaces the costlier conventional power producers and shifts the supply curve to the right and thereby lowers the day-ahead price.

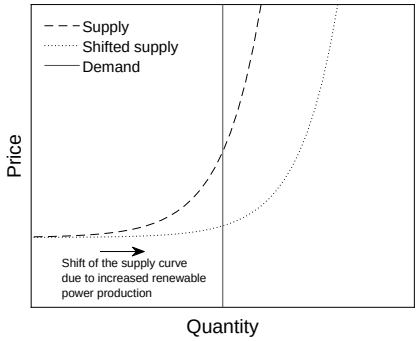


Fig. 1: Illustration of the price formation of day-ahead markets in Europe, and the impact of renewable power production.

On the other hand, the intermittent nature of renewables demands a backbone of programmable electricity sources that can deliver electricity when e.g. the wind does not blow or the sun does not shine. Referring to Fig. 1, a left shift instead of the shown right shift of the supply curve would increase the day-ahead price. To put it another way, a high amount of electricity produced by renewables puts a downward pressure on the day-ahead prices, while a low amount puts an upward pressure on the day-ahead prices. The uncertainty of the amount of produced electricity from renewables, also known as volumetric risk, translates through the day-ahead price formation to volatility in the day-ahead price. The increased integration of renewables does therefore not only impact the renewables themselves, but all market participants.

In Europe, Germany has the highest installed wind power capacity, reaching 59 GW in 2018 (Wind Europe (2019)). Incentivized by this high amount of wind power capacity, the so-called wind power futures contracts were introduced in Germany on both NASDAQ in 2015 and the European Energy

Exchange in 2016. The purpose of the wind power futures is to facilitate the hedging of the volumetric risk implied by the wind power production; hence, the underlying of the wind power futures is an index describing the overall utilization of the German wind power capacity. Loosely speaking, conventional power producers and wind power producers represent the different sides of the wind power futures market. Conventional power producers are concerned with high-wind scenarios, cf. Fig. 1, whereas wind power producers are harmed by low-wind scenarios.

In **Paper A**, co-authored with Anca Pircalabu and Esben Høg, we assess the hedging benefits of wind power futures for conventional power producers represented by a gas-fired power plant located in Germany. Based on copula theory, we propose a seasonal copula mixture for the dependence structure between the clean spark spread and the underlying wind index of the wind power futures. With regard to the marginal distributions, we employ ARMA-GARCH models. Monte Carlo simulations from the model allow us to quantify the hedging effectiveness of wind power futures for gas-fired power plants, and the implication of asymmetry, tail dependence, and seasonality in the dependence structure.

The geographical location of the conventional power producer is of less importance when assessing the hedging benefits of wind power futures, since the cost of the fuel used to generate electricity is not dependent on the specific location within Germany. On the contrary, if we consider the other side of the wind power futures market represented by wind power producers, the wind used to generate electricity is very dependent on the specific location. Since the underlying of the exchange-traded wind power futures contract represents the overall utilization of the German wind power capacity, the hedging benefits for wind power producers at different geographical locations might differ significantly. We consider in **Paper B**, co-authored with Anca Pircalabu, the spatial hedging benefits of wind power futures for wind power producers located at 31 different sites in Germany. For each wind power producer, we propose an ARMA-GARCH copula model for the joint behaviour of the site-specific capacity utilization and the underlying index of the wind power futures. As in Paper A, we use Monte Carlo simulations from the 31 different models to assess the hedging benefits and the implication of the risk premium in the wind power futures market.

The non-perfect hedge implied by the exchange-traded wind power futures contract for wind power producers lead us in **Paper C**, co-authored with Fred Espen Benth and Victor Rohde, to propose tailor-made wind power futures for wind power producers. The underlying of each tailor-made wind power futures contract is the utilization of the capacity of each wind power producer, thus implying a perfect hedge of the volumetric risk. From a modeling perspective, we propose two continuous-time models based on Lévy processes for the simultaneous modeling of several wind indexes; that is, the

simultaneous modeling of the utilization of wind power capacity at different sites. Compared to Papers A and B, relying on this continuous-time framework limits in general the flexibility of the model and its ability to capture e.g. asymmetry or tail dependence in the dependence structure. On the other hand, this framework allows us to perform calculations more efficiently from a computational point of view compared to the Monte Carlo simulations in Papers A and B. Using our two multivariate models, we investigate the usefulness of tailor-made wind power futures for both the selling and buying side. More specifically, we consider the situation of an energy management company acting as counterparty for three wind power producers geographically located at different sites in Germany.

The Role of Interconnectedness

The development of the interconnectedness between member states was and is a cornerstone in achieving one pan-European electricity market. According to the European Union directive 2009/72/EC,

"A secure supply of electricity is of vital importance for the development of European society, the implementation of a sustainable climate change policy, and the fostering of competitiveness within the internal market. To that end, cross-border interconnections should be further developed in order to secure the supply of all energy sources at the most competitive prices to consumers and industry within the Community."

From a physical point of view, this corresponds to expanding the cables that allow the transferring of electricity between member states thorugh interconnectors. The so-called electricity interconnection target, formulated by the Barcelona European Council (2002), implies that for each member state the maximum possible export of electricity through interconnectors to surrounding member states should at least be 10% of the installed power production of that particular member state by 2020. While some member states have met this requirement, the drastic change in the energy production mix towards more volatile and non-programmable energy sources such as wind and solar indicates that more interconnector capacity might be necessary to benefit the most from such energy sources, as pointed out by the Commission Expert Group (2017).

Nonetheless, without proper coordination between member states and a joint plan for the use of the interconnectors, increased interconnectedness is futile. Historically, explicit auctions for PTRs have been applied to allow third parties access to the interconnectors (see Füss *et al.* (2015) for more information). Here explicit refers to the fact that the allocation of the right to transfer electricity for a certain cross-border happens independently of the trading of electricity. Market participants that have required PTRs have

to schedule the flows of electricity themselves, before actually knowing the day-ahead prices. As a result, sub-optimal flows of electricity occur. To exemplify, according to Danish Energy Agency (2018), the "*interconnectors were only optimally utilised in 30 pct. of the time, and that the electricity even flowed in the "wrong" direction in 24 pct. of the time*" for the electricity flowing between West Denmark and Germany.

Implicit allocation of day-ahead transmission rights is a way of mitigating this. The Price Coupling of Regions (PCR) is a project of eight power exchanges[2] with the main goal of allocating day-ahead transmission rights implicitly in the European day-ahead electricity markets. One single algorithm, known as Euphemia (Pan-European Hybrid Electricity Market Integration Algorithm), takes bids and offers for delivery of day-ahead electricity at all member exchanges as input, as well as the physical restrictions of the electricity transmission grid.[3] Based on an objective of optimizing the overall social welfare, Euphemia then calculates optimal flows of electricity between the market areas and the corresponding day-ahead prices. In this way, market areas are coupled, which is why this implicit allocation also goes under the name *market coupling*. The Euphemia algorithm has been applied since February 2014, where it gradually replaced its predecessors that only took into account individual parts of the PCR area. For more information on the Euphemia algorithm, the reader is referred to NEMO Committee (2019).

As a consequence of the market coupling, there is no need for market participants to schedule day-ahead PTRs and potentially flowing electricity inefficiently. Regarding the nomination of long-term PTRs, if they are sold under the UIOSI principle, the incentive to nominate them is rather low. As reported in ENTSOE (2012), the combination of long-term PTRs with UIOSI principle and the market coupling implies that the amount of time that long-term PTRs are nominated has decreased significantly. Hence, from a financial point of view, there is no difference between an FTR and a PTR under UIOSI principle in a market coupled environment.

The shift in the price formation in coupled interconnected markets has introduced significant changes on day-ahead electricity prices and derivatives such as transmission rights. Most prominent is the presence of exact price convergence between day-ahead prices; i.e., due to the market coupling, identical prices are observed in interconnected markets. In **Paper D**, co-authored with Fred Espen Benth, we propose a continuous-time model for day-ahead electricity prices in two interconnected markets that takes the market coupling into account. Specifically, we assume that the observed day-ahead

[2] "*EPEX SPOT, GME, HEnEx, Nord Pool, OMIE, OPCOM, OTE and TGE covering the electricity markets in Austria, Belgium, Czech Republic, Croatia, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, Ireland, Latvia, Lithuania, Luxembourg, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and UK*", according to NEMO Committee (2019).

[3] The transmission system operator (TSO) of each market area provides this information.

prices are driven by Ornstein-Uhlenbeck processes representing the partly latent domestic day-ahead prices, and a latent process dictating whether the observed day-ahead prices in the two market areas are equal or not. Due to the latent nature of the model, we rely on filtering techniques to estimate the model. The proposed estimation procedure is divided into a Kalman filtering part and a particle filtering part. Lastly, we discuss in a theoretical setting the pricing of forwards and long-term transmission rights implied by the model.

An empirical investigation of the proposed model in Paper D is the theme of **Paper E**, co-authored with Jesper Jung. Here we consider the trading of long-term FTRs (or PTRs under UIOSI principle) in interconnected coupled electricity markets. The model is assessed in a pricing and hedging context. In both cases, we consider auction prices for 22 monthly[4] transmission rights spanning the period from August 2017 to May 2019 for each considered cross-border connection. As hedging instruments, we use baseload forward contracts of varying delivery periods. The spread option formula from Margrabe (1978) is used throughout the paper as a benchmark due to its well-known properties.

[4]That is, the delivery period of the transmission right is a month.

References

Barcelona European Council (2002). Presidency conclusions, 15 and 16 March 2002.

Commission Expert Group on electricity interconnection targets (2017). Towards a sustainable and integrated Europe.

Danish Energy Agency (2018). European experiences on power markets facilitating efficient integration of renewable energy.

ENTSOE (2012). Transmission risk hedging products. *An entso-e educational paper*.

European Commission (2016). Energy prices and costs in Europe. COM(2016) 769 final, 30.11.2016.

European Union. Directive 2009/72/EC of the European Parliament and of the Council concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC. *Official Journal of the European Union*.

Füss, R., Mahringer, S., Prokopczuk, M. (2015). Electricity market coupling in Europe: status quo and future challenges. *Working paper on finance no. 2015/12, University of St. Gallen*, **14**.

Heddenhausen, M. (2007). Privatisations in Europe's liberalized electricity markets – the cases of the United Kingdom, Sweden, Germany and France. *German Institute for International and Security Affairs*.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance*, **33**, pp. 177–186.

NEMO Committee (2019). Euphemia public description. Single price coupling algorithm.

Wind Europe (2019). Wind energy in Europe in 2018. Trends and statistics.

Part II

Papers

Paper A

A Seasonal Copula Mixture for Hedging the Clean Spark Spread with Wind Power Futures

List of authors: Troels Sønderby Christensen[1,2], Anca Pircalabu[1,2], Esben Høg[1]

[1]Department of Mathematical Sciences, Aalborg University
[2]Quantitative Analytics, Centrica Energy Trading

The paper has been published in
Energy Economics Vol. 78, pp. 64–80, February 2019.

The layout has been revised.

Abstract

The recently introduced German wind power futures have brought the opportunity to address the problem of volume risk in wind power generation directly. In this paper, we study the hedging benefits of these instruments in the context of peak gas-fired power plants, by employing a strategy that allows trading in the day-ahead clean spark spread and wind power futures. To facilitate hedging decisions, we propose a seasonal copula mixture for the joint behavior of the day-ahead clean spark spread and the daily wind index. The model describes the data surprisingly well, both in terms of the marginals and the dependence structure, while being straightforward and easy to implement. Based on Monte Carlo simulations from the proposed model, the results indicate that significant benefits can be achieved by using wind power futures. Moreover, a comparison study shows that accounting for asymmetry, tail dependence, and seasonality in the dependence structure is especially important in the context of risk management.

1 Introduction

The sudden change in German energy policy that followed the Fukushima nuclear accident marked a new era for the German power market. Since the nuclear shutdown and the shift to renewables, Germany has experienced an impressive growth in both wind and solar power, and has reached a level that far exceeds the Kyōto climate obligations. This change has undoubtedly brought benefits on several fronts, however, the non-programmable nature of wind and solar electricity production has resulted in a large share of weather-dependent supply of electricity. From a financial point of view, the cash-flows from such non-programmable power plants can be incredibly volatile, not only due to price uncertainty, but also due to the uncertainty associated with the volume produced. While renewable generators are clearly affected by the uncertain volume, they are not the only ones; by market design and economics principle, the presence of renewables in the bid stack will always force conventional generators to produce less. In Germany, where the share of renewable energy is especially high, the conventional producers' competitiveness on e.g. the spot and forward markets has deteriorated, which has in turn invoked the need for far more intricate operation patterns and strategies.

In light of the advancements concerning renewables in Germany and the challenges imposed by volume risks for many different market players, the European Energy Exchange (EEX) recently introduced a financial instrument to mitigate the volume uncertainty associated with wind power generation. This instrument is referred to as a *wind power futures*, and its underlying is the German wind index. Representative agents for the sell and buy sides of wind power futures are the wind electricity producers and the conven-

tional electricity producers, respectively. On one hand, low wind scenarios are unfavorable for wind electricity producers, since they have a lowering effect on cash-flows; on the other hand, conventional generators are exposed to high wind scenarios, since a large share of wind power in the electrical grid displaces the costlier sources.

In this paper, we study the risk-reducing benefits of wind power futures in the context of conventional generators that operate in the day-ahead market whenever profitable. As a representative agent for the conventional generator, we consider the case of a peak gas-fired power plant whose profit per unit of electricity produced is measured in terms of the day-ahead clean spark spread. Since the dependence between the day-ahead clean spark spread and the wind index is essential for assessing the benefits of wind power futures, the contribution of this paper is twofold.

First, we propose a seasonal copula mixture to model the joint behavior of the day-ahead clean spark spread and the daily wind index. The model is fitted to four years of German data, and captures the marginal behavior of the individual variables and also the seasonality in the dependence between the variables very well. Second, we employ the proposed seasonal copula mixture to facilitate hedging decisions and showcase the effectiveness of wind power futures. To highlight the benefits of the seasonal copula mixture, we perform a study where the proposed model is compared against alternative models.

Owing to the recent introduction of the German wind power futures, the related literature is very scarce. The first related study is that of Gersema and Wozabal (2017), where the authors focus mainly on the pricing of wind power futures and explaining risk premia, for which an equilibrium pricing model is proposed. Also concentrating on the pricing aspect is the work of Benth and Pircalabu (2018), who apply a no-arbitrage approach to the pricing of wind power futures, and obtain results concerning the sign of risk premia that support the conclusions drawn in Gersema and Wozabal (2017). In contrast to the two existing studies, which focus mainly on pricing and less on hedging and risk management, we take a simplistic approach to pricing but study in detail aspects related to the risk-reducing ability of wind power futures. Nevertheless, we acknowledge that some of the results in Gersema and Wozabal (2017) and Benth and Pircalabu (2018) are very relevant in the context of the present study, and they shall thus be included in our discussion.

Turning to applications of copulas in energy markets, we mention that these models have gained substantial interest over the past years and have become a popular tool to model the non-linear dependence between different commodities. Some examples concerning applications of bivariate copulas are Börger *et al.* (2009), Benth and Kettler (2011), Grothe and Schneiders (2011), Avdulaj and Barunikl (2015), and Elberg and Hagspiel (2015). For ap-

plications beyond bivariate copulas, we mention the study of Pircalabu and Jung (2017), and that of Aepli *et al.* (2017). The present paper contributes to this stream of literature in terms of the application, which to the best of our knowledge has not yet been considered, and also in terms of modeling approach, by proposing an extension that deals with seasonality in the dependence structure.

The remaining of this paper is structured as follows: In Section 2, we introduce the data and elaborate on the construction of the variables. In Section 3, we describe the modeling framework and report estimation results. Section 4 introduces the seasonal copula mixture model and provides evidence for its quality of fit. In Section 5, we employ the proposed model to study the benefits of wind power futures, and perform various comparison studies. Section 6 concludes.

2 Background and Data

To investigate the benefits of wind power futures for a gas-fired power plant (GFPP), two data components are of interest in the analysis performed in this paper: The day-ahead clean spark spread and the daily wind index. In this section, we address each of these in turn, commenting on their construction.

2.1 Clean Spark Spread

As an indicator for the profit per unit of electricity generated by a GFPP, we consider the day-ahead clean spark spread (CSS). This measure depends on electricity, gas, and emission prices, and also on the heat rate and the emission factor. The heat rate represents the required number of natural gas MWhs to produce one MWh electricity, i.e., the efficiency at which the GFPP transforms gas to electricity. Further, the emission factor represents the number of tons of CO_2 emitted by producing one MWh electricity.

With GFPPs being mainly peak-operated power plants—that is, power plants dispatching during the peak hours between 8 AM to 8 PM on weekdays and non-holidays—we consider the peak electricity price. Specifically, let S_t^E denote the day-ahead peak load electricity price, S_t^G the day-ahead gas price, and S_t^C the day-ahead emission price, with the subscript t indicating time measured in days. Further, let h be the heat rate and e the emission factor. We define the day-ahead CSS on day t as

$$CSS_t = S_t^E - hS_t^G - eS_t^C, \tag{A.1}$$

where S_t^E and S_t^G are measured in EUR/MWh, and S_t^C is measured in EUR/tCO_2.

Data Preparation for the Clean Spark Spread

To construct a time series for the day-ahead CSS, we consider the following time series:

- S_t^E: The German electricity price, which is computed as the average of all hourly electricity prices between 8 AM to 8 PM on weekdays and non-holidays. The source of this data is EEX.
- S_t^G: The day-ahead gas price for NetConnect Germany (NCG), which corresponds to the closing price. The source of this data is EEX.
- S_t^C: The EU Allowance unit of one tonne of CO_2 (EUA) phase 3 daily futures price. This data is collected from the Intercontinental Exchange, and represents the closing price. For more information regarding the EU emissions trading system, we refer the interested reader to European Commission (2017).

All time series above span 1030 observations in the period from 3 January 2013 to 30 December 2016, and cover weekdays that are non-holidays. To provide a sense of the data, we plot in Fig. A.1 the time series corresponding to each of the three data sources described above. Moreover, we plot the CSS obtained by applying Eq. (A.1), and using the values for h and e reported in Table A.1. These numbers are based on ICIS (2016), and shall be used in the remaining of this paper unless explicitly stated otherwise.

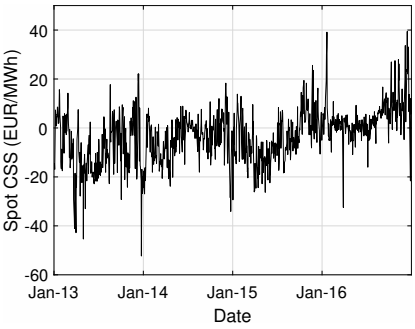
Heat rate h	Emission factor e
2.035	0.375

Table A.1: Heat rate (MWhs natural gas per MWh electricity) and emission factor (tCO_2 per MWh electricity) based on ICIS (2016). The chosen heat rate corresponds to an efficiency of 49.13%.

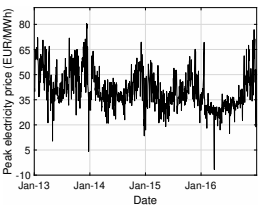
2.2 Wind Index

Since the German wind power futures (WPF) were introduced only recently, we find it relevant to provide a brief description of these products and to clarify their payoff structure. WPF contracts are written on the average wind index in Germany, and can be traded at the European Energy Exchange (EEX) and Nasdaq OMX. In this paper, we shall restrict our attention to the WPF traded at EEX.

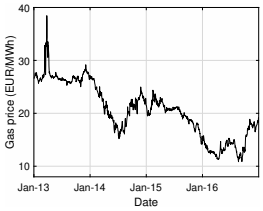
The German wind index is obtained as the ratio between the total wind power generation and the total available installed wind power capacity. Hence, the index is bounded between zero and one, and provides a measure of the German wind utilization. Currently, delivery periods for WPF



(a) CSS_t: Daily day-ahead CSS (peak load)



(b) S_t^E: German day-ahead electricity price (peak load)



(c) S_t^G: NCG gas day-ahead closing price

(d) S_t^C: EUA day-ahead closing price

Fig. A.1: Historical evolution of the daily day-ahead CSS (peak load), the German day-ahead electricity price (peak load), the NCG day-ahead gas closing price, and the EUA day-ahead closing price, from 3 January 2013 to 30 December 2016. The applied heat rate and emission factor to construct the day-ahead CSS are given in Table A.1.

correspond to weeks, months, quarters and years, and only trading the base load profile is possible. Compared to the definition of the day-ahead CSS data in Eq. (A.1), there is clearly a mismatch between delivery periods, with wind power futures hedging all hours of every day, and gas turbines generating output during peak hours. However, this reflects the present market conditions, where the volume risk of a GFPP can only be imperfectly hedged. Assuming a delivery period $[T_1, T_2]$ consisting of H hours, the payoff corresponding to a long position in one WPF contract is given by

$$R^{WPF} = H \left(\underbrace{\frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} W_t}_{=\bar{W}} - W_{t_0} \right) X, \tag{A.2}$$

where $W_t \in [0,1]$ is the daily wind index, $\bar{W} \in [0,1]$ is the realized average wind index over the delivery period, and $W_{t_0} \in [0,1]$ can be thought of as the "futures price", i.e., the index set at t_0 when entering the contract. Further, X is a pre-specified constant tick size which is used to convert the index differences into monetary measures. According to EEX, $X = 100$ EUR. We see from Eq. (A.2) that a short position in WPF will generate a profit in low-wind scenarios, making it a useful hedging instrument for the wind power producer. Conversely, a long position will generate a profit in the high-wind scenarios, to which the GFPPs are exposed.

Data Preparation for the Wind Index

The index that a WPF contract is settled against is externally provided by EuroWind. Since trading in WPF started only recently, the amount of data available on the spot wind index provided by EuroWind is limited. To obtain a longer time series, we consider instead a proxy wind index constructed using the wind power production in Germany on a daily basis, and monthly recordings of the German installed wind power capacity, which are updated at the start of each month. The wind power production data is collected from the four different transmission system operators in Germany, and the source of the installed capacity data is PointConnect.

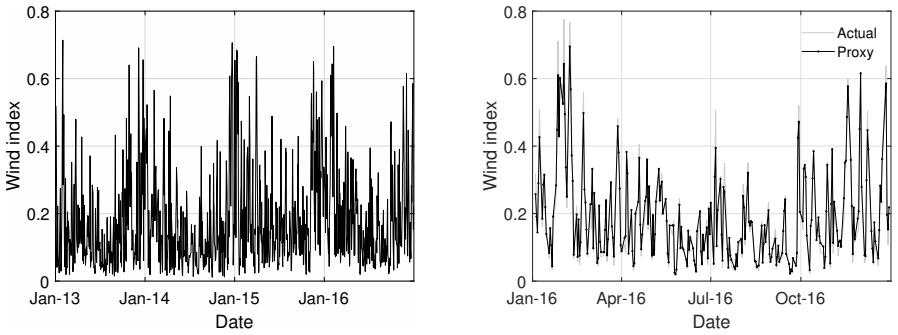
Specifically, the daily German wind index is constructed as

$$W_t = \text{Daily wind index} = \frac{\text{Daily wind power generation (MWh)}}{H_t \cdot \text{Installed capacity (MW)}},$$

where H_t denotes the number of hours in day t, and the installed capacity on a daily basis is obtained by linear interpolation. In order to unify the length of the day-ahead CSS and the wind index, we omit weekends and holidays for the wind index data. Hence, the constructed index spans the period from

3 January 2013 to 30 December 2016, a total of 1030 observations, and is plotted in Fig. A.2(a).

To provide some evidence for how the constructed wind index matches the true settlement data, we plot in Fig. A.2(b) our proxy together with the one year of actual data from EuroWind that we have available. The time series plot reveals an acceptable resemblance, and to provide a quantitative indication, we compute the mean absolute error to 0.020.



(a) Wind index (b) Comparison with EuroWind data

Fig. A.2: Historical evolution of the German wind index (W_t) on a daily basis from 3 January 2013 to 30 December 2016, and a comparison of W_t with the actual wind index provided by EuroWind for the year 2016.

3 Model Construction and Fit

To model the joint behavior of the day-ahead CSS and the daily wind index (henceforth referred to as simply CSS and wind index, respectively), we consider copula models. Restricting our presentation to the two-dimensional case, a copula is the joint distribution of the random variables U_1 and U_2, where each variable is marginally uniformly distributed as Unif(0,1). Since our data exhibits seasonality and autocorrelation, we wish to filter out these effects before applying the copula. Therefore, we are here considering the conditional copula.

Let $F(\cdot|\mathcal{F}_{t-1})$ denote the conditional joint distribution function of the random vector $\boldsymbol{Y}_t = (Y_{1t}, Y_{2t})$, and let $F_1(\cdot|\mathcal{F}_{t-1})$ and $F_2(\cdot|\mathcal{F}_{t-1})$ denote the conditional continuous marginal distribution functions of Y_{1t} and Y_{2t}, respectively. Then, according to Sklar's theorem [Sklar (1959)] for conditional distributions, there exists a unique copula C such that F can be decomposed as

$$F(y_{1t}, y_{2t}|\mathcal{F}_{t-1}) = C(F_1(y_{1t}|\mathcal{F}_{t-1}), F_2(y_{2t}|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}). \tag{A.3}$$

The converse also holds, meaning that given two univariate distributions F_1, F_2 and a copula C, F as defined in Eq. (A.3) is the joint distribution with margins F_1, F_2. Thus, Sklar's theorem not only provides a way of *decomposing* a joint distribution function, but also a way of *composing* it given marginal distributions and a copula, both of which are very useful in practical applications. For the proof of Sklar's theorem for conditional distributions, we refer to Patton (2006(a)).

Recalling the probability integral transform, we note that $U_{it} := F_i(Y_{it}|\mathcal{F}_{t-1}) \sim \text{Unif}(0,1)$, for $i = 1, 2$. Differentiating both sides of Eq. (A.3) with respect to (y_{1t}, y_{2t}) thus yields

$$f(y_{1t}, y_{2t}|\mathcal{F}_{t-1}) = c\left(u_{1t}, u_{1t}|\mathcal{F}_{t-1}\right) \cdot f_1(y_{1t}|\mathcal{F}_{t-1}) \cdot f_2(y_{2t}|\mathcal{F}_{t-1}), \tag{A.4}$$

where f denotes the joint density function, c is the copula density, and f_1, f_2 denote marginal density functions.

In our context, copula models are advantageous for various reasons: First, being able to capture dependence beyond the linear correlation can be of utmost importance when illustrating the hedging benefits of WPF, and this can be achieved with copulas. Second, we can separate the treatment of the dependence structure from that of the marginal behavior of the individual variables (cf. Eq. (A.4)), since the dependence structure is fully contained in the copula. Third, selecting one type of marginal distribution for the first variable does not restrict our choice of marginal distribution for the second variable.

Turning to the estimation of the model parameters, we let T denote the sample size, $\boldsymbol{\theta}_c$ the copula parameters, and $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ the parameters of the marginal models. From Eq. (A.4) it follows that the log-likelihood function is

$$\begin{aligned}\log \mathcal{L} = &\sum_{t=1}^{T} \log c\left(u_{1t}, u_{1t}|\mathcal{F}_{t-1}, \boldsymbol{\theta_c}\right) + \sum_{t=1}^{T} \log f_1(y_{1t}|\mathcal{F}_{t-1}, \boldsymbol{\theta_1}) \\ &+ \sum_{t=1}^{T} \log f_2(y_{2t}|\mathcal{F}_{t-1}, \boldsymbol{\theta_2}).\end{aligned}$$

Here, we consider multi-stage maximum likelihood (MSML) estimation. This provides a far less complicated estimation procedure relative to one-stage MLE. Moreover, the studies of Joe (2005) and Patton (2006(b)) suggest that the efficiency loss is not substantial.

In the following two sections, we present in detail the marginal models and the constant copulas considered in this paper, and provide empirical evidence for the fit of these models to our data.

3.1 Marginal Models

Since both the CSS and the wind index exhibit seasonality, we start the marginal treatment of the individual variables by applying suitable seasonal functions to remove the deterministic seasonal component.

For the CSS, we consider the seasonal function

$$f_t = a_1 + b_1 t + c_1 \sin(2\pi t/K) + c_2 \cos(2\pi t/K),$$

where a_1 is a constant, b_1 is the trend coefficient, and c_1 and c_2 are coefficients for the annual cycle. We have on average approximately $K = 258$ observations per year.

Not surprisingly, the seasonality function for the CSS resembles a seasonality function that would typically be considered for the day-ahead electricity price (see e.g. Haldrup and Nielsen (2006), Benth and Šaltytė Benth (2011), and Härdle and López Cabrera (2012)). This resemblance is caused by the magnitude of the electricity price compared to the gas and emission price, cf. Fig. A.1, causing the former to have the dominant effect. Aside from electricity prices usually exhibiting a yearly seasonality, a strong within-week seasonality is also often observed. However, with the exclusion of weekends from our data, adding a term that addresses the weekly seasonality (e.g. day-of-week dummies) is unnecessary. Furthermore, adding more trigonometric terms (based on the periodogram) does not improve the fit of the seasonal function substantially.

Turning to the wind index, recall that this series is bounded between 0 and 1, cf. Fig. A.2. Following Pircalabu and Jung (2017), we apply the logit-transform to the wind index[1], and consider the following seasonal function for the logit wind index (LWI):

$$f_t = a_1 + c_1 \sin(2\pi t/K) + c_2 \cos(2\pi t/K),$$

which is motivated by the prominent annual cycles we observe in the sample autocorrelation of the LWI. Also here, different meaningful extensions of the seasonal function were experimented with, without yielding a significant improvement.

The seasonality functions are fitted to the data by ordinary least squares, and Table A.2 summarizes the results obtained for the CSS and the LWI.

Next, we apply ARMA-GARCH filters to the deseasonalized data. Given

[1]The logit function is given by $\text{logit}(x) = \log(x) - \log(1-x)$.

	$\hat{a_1}$	$\hat{b_1}$	$\hat{c_1}$	$\hat{c_2}$
CSS	-10.131	0.014	-3.611	2.990
	(0.579)	(0.001)	(0.411)	(0.404)
LWI	-1.828	-	0.066	0.472
	(0.030)	-	(0.042)	(0.043)

Table A.2: OLS estimates for parameters of the seasonal functions for the CSS and the LWI. Standard errors are reported in parenthesis, and are based on a naive OLS calculation.

a time series of data y_t, an ARMA(p,q)–GARCH(h,k) model is defined by

$$\begin{aligned} y_t &= \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t, \\ \varepsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^{h} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{k} \beta_j \sigma_{t-j}^2, \end{aligned}$$

where $\eta_t \sim iid\ \mathrm{N}(0,1)$. It was Engle (1982) who introduced the ARCH model, and later Bollerslev (1986) who extended the variance equation to include lagged values of σ_t^2. For a review of ARMA and GARCH models, see for example Shumway and Stoffer (2007). In the following, we denote by η_t^{CSS} and η_t^{LWI} the standardized residuals resulting from applying the ARMA–GARCH models to the CSS and the LWI, respectively.

Model selection is based on the Bayesian Information Criterion (BIC), and we consider ARMA(p,q)–GARCH(h,k) models for all possible combinations of p,q,h and k, for $p = 0,\ldots,7$, $q = 0,\ldots,7$, $h = 0,1,2$, and $k = 0,1,2$. The optimal order of the models and the corresponding estimated parameters are reported in Table A.3.

Considering the goodness-of-fit of the normal distribution, we find a satisfactory fit in the case of $\hat{\eta}_t^{LWI}$, cf. Figs. A.3(e) and A.3(f). This is however not the case for $\hat{\eta}_t^{CSS}$. Consequently, we relax the normality assumption for the CSS, and consider instead the normal-inverse Gaussian (NIG) distribution. The probability density function of the NIG distribution is given by

$$g(x|\alpha,\beta,\mu,\delta) = \frac{\alpha\delta G_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi\sqrt{\delta^2+(x-\mu)^2}} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)},$$

where

$$G_1(x) = \frac{1}{2}\int_0^\infty e^{-\frac{1}{2}x(t+t^{-1})}dt$$

is the modified Bessel function of third kind and index 1. The NIG distribution is a popular choice in the financial literature (for some examples, see

	CSS	**LWI**
Model	ARMA(2,1) – GARCH(1,1)	ARMA(1,1)
	Conditional mean	
AR1 $\hat{\phi}_1$	1.313 (0.077)	0.360 (0.055)
AR2 $\hat{\phi}_2$	-0.360 (0.058)	-
MA1 $\hat{\theta}_1$	-0.818 (0.064)	0.208 (0.055)
Variance $\hat{\sigma^2}$	-	0.677 (0.030)
	Conditional variance	
Constant $\hat{\omega}$	3.611 (0.636)	-
ARCH $\hat{\alpha_1}$	0.110 (0.022)	-
GARCH $\hat{\beta_1}$	0.827 (0.027)	-

Table A.3: Type and order of marginal models, parameter estimates and corresponding standard errors in parenthesis.

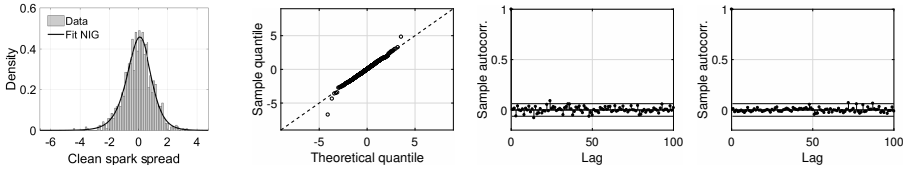
Barndorff-Nielsen (1997a), Rydberg (1997), Barndorff-Nielsen (1997b), and Jensen and Lunde (2001)), and is also often able to provide a good description of commodity data, see e.g. Benth and Šaltytė Benth (2004) and Benth and Kettler (2011). The NIG distribution is fitted to the residuals from the ARMA(2,1)–GARCH(1,1) model cf. Table A.3 by maximum likelihood, and the parameter estimates are reported in Table A.4. As it appears from the histogram and quantile plots displayed in Figs. A.3(a) and A.3(b), the NIG distribution provides a satisfactory fit to the CSS data.

$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\delta}$
1.584 (0.307)	-0.189 (0.144)	0.189 (0.126)	1.534 (0.281)

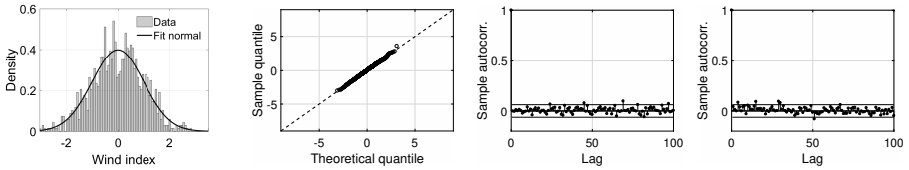
Table A.4: Maximum likelihood estimates obtained by fitting the NIG distribution to $\hat{\eta}_t^{CSS}$. Corresponding standard errors are given in parenthesis.

To provide further evidence for the appropriateness of the chosen marginal distributions, we perform the Kolmogorov-Smirnov (K-S) and the Cramer-von Mises (CvM) goodness-of-fit tests. To obtain critical values for the tests, we employ the simulation-based method described in detail in Patton (2013). In the CSS case we obtain p-values of 0.627 and 0.785 for the K-S and CvM test, respectively, and can thus not reject the null that the NIG distribution is well-specified. This is also the conclusion in the LWI case, where we test the goodness-of-fit of the normal distribution. Here, the resulting p-values are 0.746 and 0.915 for the K-S and CvM test, respectively.

Aside from providing evidence for the goodness-of-fit of the marginal distributions, the sample autocorrelations provided in Fig. A.3 suggest that no considerable serial dependence is left in the conditional mean and variance, for either variable. Having verified that the models proposed here are suit-



(a) Histogram for $\hat{\eta}_t^{CSS}$ (b) Quantile plot for $\hat{\eta}_t^{CSS}$ (c) Acf for $\hat{\eta}_t^{CSS}$ (d) Acf for squared $\hat{\eta}_t^{CSS}$



(e) Histogram for $\hat{\eta}_t^{LWI}$ (f) Quantile plot for $\hat{\eta}_t^{LWI}$ (g) Acf for $\hat{\eta}_t^{LWI}$ (h) Acf for squared $\hat{\eta}_t^{LWI}$

Fig. A.3: Diagnostics for the standardized residuals $\hat{\eta}_t^{CSS}$ (first row) and $\hat{\eta}_t^{LWI}$ (second row).

able for describing the marginal behavior of the CSS and the LWI, we proceed in the next section to the modeling of the dependence structure.

3.2 Constant Copula Models

Let F^{NIG} and F^N denote the cumulative distribution functions for the NIG and standard normal distribution, respectively. To obtain the approximately uniforms that are the input variables to the copula function, we apply the probability integral transform, i.e.,

$$\hat{u}_t^{CSS} = F^{NIG}(\hat{\eta}_t^{CSS}|\mathcal{F}_{t-1}, \hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\delta}),$$
$$\hat{u}_t^{LWI} = F^{N}(\hat{\eta}_t^{LWI}|\mathcal{F}_{t-1}),$$

for $t = 1, \ldots, T$. In Fig. A.4 we plot the resulting probability integral transforms against each other, revealing that the variables are negatively related. This finding is not surprising considering the negative dependence between electricity prices and the wind index. A high wind penetration in the electricity grid puts a downward pressure on day-ahead electricity prices owing to the process of day-ahead price formation, which prioritizes cheap electricity producers. With everything else being equal, this lowering effect on the electricity price is then reflected in the CSS, which is also lowered, cf. Eq. (A.1). Similar arguments apply to the case of a low wind scenario, where electricity prices are typically pushed upwards.

The dependence structure seems to be slightly asymmetric, with the north west corner of Fig. A.4 exhibiting more concentration and being sharper in

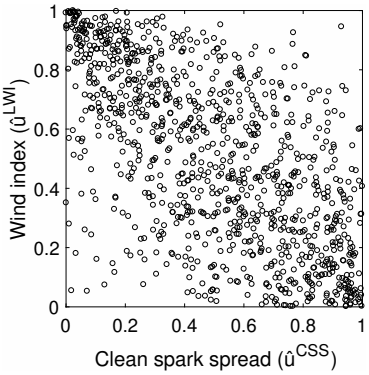


Fig. A.4: Empirical copula density.

shape compared to the south east corner. That is, there seems to be more probability of observing the combination of extremely high wind / extremely low CSS than the reverse. Since non-zero dependence in extreme events could have notable implications for the benefits of WPF, being able to capture such behavior in a model for the dependence structure must be considered. Luckily in the context of copulas, such extreme events can be easily captured by considering certain copula families that allow for non-zero *tail dependence*. More specifically, the lower and upper tail dependence can be defined as

$$\Lambda^l = \lim_{q \to 0^+} P(u_t^{CSS} \leq q | u_t^{LWI} \leq q)$$
$$\Lambda^u = \lim_{q \to 1^-} P(u_t^{CSS} > q | u_t^{LWI} > q),$$

where q denotes the quantile. Clearly, since our data is characterized by negative dependence, computing Λ^l and Λ^u as defined above is not meaningful. This can however be resolved by performing suitable rotations of the data, which shall be discussed in more detail shortly.

To investigate which copula best describes the dependence structure illustrated in Fig. A.4, we consider first the following standard copulas, which are often employed in the related literature: Gaussian, Gumbel, rotated Gumbel (RGumbel), Clayton, rotated Clayton (RClayton), Frank, symmetrized Joe-Clayton, and Student t.[2] These copula models cover a wide range of dependency structures, with some models being able to capture asymmetric dependence, and also upper and lower tail dependence, i.e., a non-zero probability of extreme events happening simultaneously. In the interest of brevity, we shall not go into detail with the properties of each copula model here, and refer instead to McNeil *et al.* (2005), Nelsen (1999), and Patton (2006(a)) for a comprehensive description.

[2]By rotated, we mean a 180 degree rotation of the data.

To allow for further flexibility compared to the standard copulas enumerated above, we also consider copula mixtures. As in e.g. Rodriguez (2007) and Dias and Embrechts (2009), for a given t we mix copula a, having copula density $c^a(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a)$, with copula b, having copula density $c^b(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^b)$, by using a mixing parameter $0 < \lambda < 1$ and the following form:

$$c^m(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b, \lambda) = \lambda c^a(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a) + (1-\lambda)c^b(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^b). \tag{A.5}$$

As expected, a mixture copula inherits characteristics from its mixing components. In the following proposition, we present an especially useful result relating to the tail dependence of a mixture copula, which we shall use shortly. Notice that we omit the conditioning to ease the notation.

Proposition 1
Let $U_i \sim Unif(0,1)$ for $i = 1, 2$, and let C^m denote the bivariate copula of (U_1, U_2). Further assume C^m is given as the mixture

$$C^m(u_1, u_2) = \lambda C^a(u_1, u_2) + (1-\lambda)C^b(u_1, u_2),$$

where C^a and C^b are two bivariate copulas, and $0 < \lambda < 1$. Then, the lower tail dependence Λ^l and the upper tail dependence Λ^u for the mixture C^m are given as

$$\Lambda^l = \lambda\Lambda^{l,a} + (1-\lambda)\Lambda^{l,b},$$

and

$$\Lambda^u = \lambda\Lambda^{u,a} + (1-\lambda)\Lambda^{u,b},$$

where $\Lambda^{l,a}$, $\Lambda^{l,b}$, $\Lambda^{u,a}$, and $\Lambda^{u,b}$ are the respective tail dependence measures for C^a and C^b.

Proof. See A.1. □

Moving on to the estimation aspect, we let $c(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta})$ denote the conditional copula density with parameter vector $\boldsymbol{\theta} \in \mathbb{R}^l$, where $l \in \mathbb{N}$ is the number of parameters in the copula. For each copula model, we obtain an estimate for $\boldsymbol{\theta}$ by maximizing the copula log-likelihood, i.e.,

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{t=1}^{T} \log c(\hat{u}_t^{CSS}, \hat{u}_t^{LWI}|\mathcal{F}_{t-1}, \boldsymbol{\theta}). \tag{A.6}$$

We note that it is only the Gaussian and Student t copulas that allow for negative dependence. To fit the remaining copulas to our data, we perform suitable rotations of the data. Specifically, we rotate around the LWI variable for the case of the Gumbel, Clayton, Frank, and symmetrized Joe-Clayton

copulas, and consider thus the pair $(\hat{u}_t^{CSS}, 1 - \hat{u}_t^{LWI})$ as input to Eq. (A.6) for these models. Regarding tail dependence, the rotation of data implies that lower tail dependence for the estimated copulas corresponds to high wind index / low CSS scenarios (north west corner of Fig. A.4), whereas upper tail dependence for the estimated copulas corresponds to low wind index / high CSS scenarios (south east corner of Fig. A.4). To fit the RGumbel and the RClayton, we note that a further 180 degree rotation of the pair $(\hat{u}_t^{CSS}, 1 - \hat{u}_t^{LWI})$ is performed.

In Table A.5, we report the estimation results for all standard copula models and three selected mixtures. Other copula mixtures aside from those reported in Table A.5 were considered, but we found no increase in performance. As a model selection criterion, we employ the Akaike Information Criterion (AIC). According to the AIC, the preferred model is the mix of Frank and RGumbel (hereafter denoted FRG copula), confirming the presence of slight asymmetry in the dependence structure illustrated in Fig. A.4.

Considering the FRG copula in more detail, its first mixing component, the Frank copula, imposes symmetric dependence and a zero tail dependence. Its second mixing component, the RGumbel, imposes an asymmetric dependence structure, with zero upper tail dependence and lower tail dependence given by

$$\Lambda^{l,RG} = 2 - 2^{1/\theta^{RG}},$$

where θ^{RG} is the parameter for the RGumbel copula. Recalling Prop. 1, we thus have that the upper and lower tail dependence for the FRG copula are

$$\begin{aligned} \Lambda^{u,FRG} &= 0, \\ \Lambda^{l,FRG} &= (1 - \lambda)\Lambda^{l,RG}. \end{aligned} \tag{A.7}$$

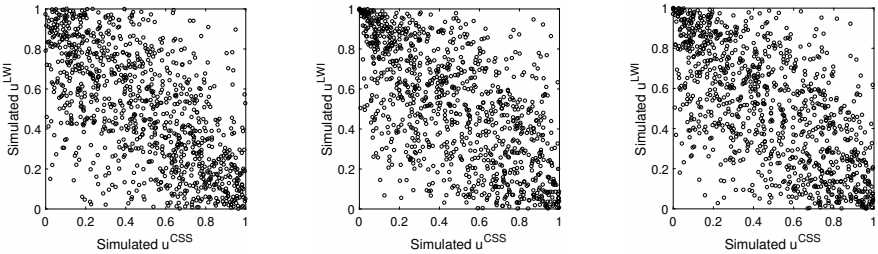
The fit produced by the FRG translates into a tail dependence coefficient of approximately 0.359 when considering the north west corner of Fig. A.4, and hence a rather high probability of extremely high wind index / low CSS happening simultaneously. To illustrate the shape of the FRG copula and how it deviates from the shapes of the individual copulas in the mixture, we plot in Fig. A.5 simulations from the fitted Frank, RGumbel and FRG copulas. The simulations reveal that while the fitted Frank copula is too symmetric and the fitted RGumbel is too asymmetric compared to the observed dependence in Fig. A.4, the fitted FRG mixture is able to dampen the individual effects, hence providing a better resemblance to the observed dependence structure.

4 Time-Varying Dependence

Up until this point, we have assumed a static model for the dependence structure, which is seldom a realistic representation. Natural follow-up ques-

Copula model		**Param. (s.e.)**	**log $\mathcal{L}_c$**	**AIC**
Gaussian	$\hat{\rho}$	-0.636 (0.019)	266.503	-531.006
Gumbel*	$\hat{\theta}$	1.675 (0.049)	226.083	-450.165
RGumbel*	$\hat{\theta}$	1.773 (0.053)	277.829	-553.657
Clayton*	$\hat{\theta}$	1.233 (0.083)	245.947	-489.894
RClayton*	$\hat{\theta}$	0.911 (0.071)	167.862	-333.724
Frank*	$\hat{\theta}$	5.029 (0.247)	267.805	-533.611
Sym. Joe-Clayton*	$\hat{\Lambda}^u$	0.274 (0.048)	265.777	-527.553
	$\hat{\Lambda}^l$	0.539 (0.026)		
Student t	$\hat{\rho}$	-0.646 (0.020)	274.596	-545.192
	$\hat{v}$	9.873 (15.008)		
Mix of Gumbel and	$\hat{\theta_1}$	1.964 (0.500)	285.060	-564.120
RGumbel*	$\hat{\theta_2}$	1.797 (0.101)		
	$\hat{\lambda}$	0.219 (0.081)		
Mix of Frank and	$\hat{\theta_1}$	4.552 (1.309)	286.419	-566.837
RGumbel*	$\hat{\theta_2}$	1.920 (0.169)		
	$\hat{\lambda}$	0.365 (0.104)		
Mix of Gaussian and	$\hat{\rho}$	-0.494 (0.092)	285.152	-564.304
RGumbel*	$\hat{\theta}$	2.095 (0.234)		
	$\hat{\lambda}$	0.357 (0.130)		

Table A.5: Estimation results for 11 selected copula models. The maximized value of the copula log-likelihood is denoted log $\mathcal{L}_c$. For the functional forms of the considered copulas and other characteristics, we refer to McNeil *et al.* (2005), Nelsen (1999) and Patton (2006(a)). A copula marked by an asterisk has been estimated using a suitable rotation of the data. Standard errors are based on 999 simulations.



(a) Frank (b) RGumbel (c) FRG

Fig. A.5: T simulations from the fitted Frank, RGumbel and FRG copulas, cf. Table A.5.

tions are therefore related to the presence and type of time variation in the dependence. In this section, we consider these questions in more detail.

To investigate the time-varying aspect we consider Spearman's ρ, which is a measure of concordance. In terms of a bivariate copula C, Spearman's ρ can be expressed as (see e.g. McNeil *et al.* (2005))

$$\rho = 12 \int_0^1 \int_0^1 C(u_1, u_2) \mathrm{d}u_1 \mathrm{d}u_2 - 3. \tag{A.8}$$

We compute Spearman's ρ between $\hat{u}_t^{CSS}$ and $\hat{u}_t^{LWI}$ based on a rolling window of 60 days. Fig. A.6 displays the results, and reveals a strong seasonal pattern in the dependence structure. According to Fig. A.6, the dependence is strongest around winter and weakest around summer. A possible explanation for this behaviour relates to the power generation mix in Germany and the import/export conditions. During winter, the increased wind power production has a direct lowering effect on the daily electricity price due to the mechanism of day-ahead electricity price formation. As argued in Section 3.2, this lowers the CSS. During summer, the lower wind power production does not have the same direct effect on the daily electricity price. If that were the case, prices should increase. The high photovoltaic production during peak periods combined with the high likelihood of being able to import cheap nuclear power from France prevents however prices from increasing. Consequently, this weakens the dependence between the wind power production and the CSS during the summer months.

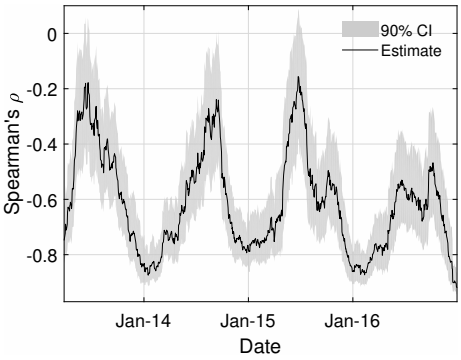


Fig. A.6: Spearman's ρ between $\hat{u}_t^{CSS}$ and $\hat{u}_t^{LWI}$ based on a 60-days rolling window. The confidence interval is based on 999 bootstraps. Note that the date corresponding to each estimate refers to the last day in the 60-days period.

In light of these findings we consider next extending the static copula mixture, such that the yearly seasonality in the dependence measured by Spearman's ρ in Fig. A.6 can be accounted for.

4.1 A Seasonal Copula Mixture

Since the FRG copula is the best performing static copula cf. Table A.5, we shall restrict our attention to this particular model in order to fix ideas. First, let us state a general result concerning Spearman's ρ for copula mixtures, which is particularly useful in our modeling context. Again here, we omit the conditioning for notational convenience.

Proposition 2
Let $U_i \sim$ Unif(0,1) for $i = 1, 2$, and let C^m denote the bivariate copula of (U_1, U_2). Further, suppose C^m is given as

$$C^m(u_1, u_2) = \lambda C^a(u_1, u_2) + (1 - \lambda) C^b(u_1, u_2)$$

for two copulas C^a and C^b, and mixing parameter $0 < \lambda < 1$. Then Spearman's ρ implied by C^m can be expressed as

$$\rho^m = \lambda \rho^a + (1 - \lambda) \rho^b, \tag{A.9}$$

where ρ^a is Spearman's ρ corresponding to copula C^a, and ρ^b is Spearman's ρ corresponding to copula C^b.

Proof. See A.2. □

It follows from Prop. 2 that Spearman's ρ for the copula mixture is simply a linear combination of the individual Spearman's ρ's corresponding to the copulas comprised in the mixture. Thus, introducing time variation in ρ^a and ρ^b translates into time variation in ρ^m. Further, it is relatively easy to compute Spearman's ρ, even for copulas where no explicit relation between Spearman's ρ and the copula parameter is available. Considering the FRG copula, the relationship between the copula parameter and Spearman's ρ for both the Frank and RGumbel copula is shown in Fig. A.7. By letting the superscripts F and RG indicate their link to the particular copula, we note that Spearman's ρ is monotonically increasing as a function of the corresponding copula parameter, θ^F and θ^{RG}. Therefore, specifying time variation for ρ^F and/or ρ^{RG} will also uniquely determine values of θ^F and θ^{RG}. If we instead were to introduce time variation directly in Spearman's ρ for the FRG copula, we would not be able to identify θ^F and θ^{RG}.

Based on the discussion above and motivated by the pronounced yearly cycle in Fig. A.6, we propose the following extension to the static FRG model. Specifically, we introduce a yearly cycle in Spearman's ρ corresponding to the RGumbel copula, i.e.,

$$\rho_t^{RG} = a^{RG} + b^{RG} \sin(2\pi t/K) + c^{RG} \cos(2\pi t/K), \tag{A.10}$$

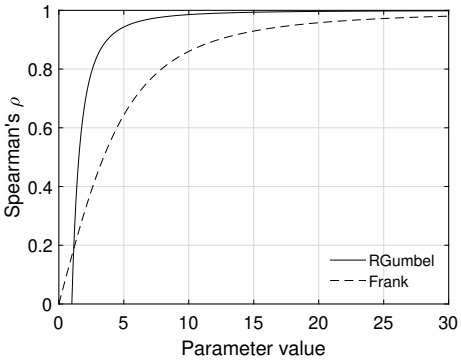


Fig. A.7: Spearman's ρ as a function of parameter value for both the Frank and RGumbel copula.

where a^{RG}, b^{RG}, and c^{RG} are constant coefficients, and $K = 258$ as was the case with the seasonal functions in Section 3.1. Regarding the Frank contribution in the FRG copula, we keep the corresponding Spearman's ρ static. Consequently, the evolution equation for the overall Spearman's ρ implied by the *seasonal* FRG copula is

$$\rho_t^{SFRG} = \lambda^{SFRG}\rho^F + (1 - \lambda^{SFRG})\rho_t^{RG}. \tag{A.11}$$

Given the seasonal specification in Eqs. (A.10)-(A.11), the model is estimated by maximizing the FRG copula loglikelihood. The estimation results are given in Table A.6, revealing a clear improvement in AIC compared to the static FRG copula.

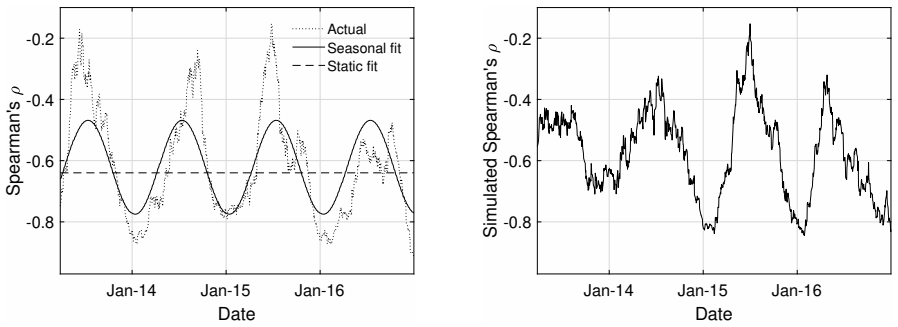
$\hat{\rho}^F$	$\hat{a}^{RG}$	$\hat{b}^{RG}$	$\hat{c}^{RG}$	λ^{SFRG}	$\log \mathcal{L}_c$	AIC
0.742	0.566	-0.132	0.208	0.316	309.894	-609.788
(0.072)	(0.041)	(0.040)	(0.043)	(0.092)		

Table A.6: Maximum likelihood estimation results for the seasonal FRG copula described in Eqs. (A.10)–(A.11). The maximized value of the copula log-likelihood is denoted $\log \mathcal{L}_c$. Standard errors are reported in parenthesis and are computed following the simulation-based procedure described in detail in Patton (2013), where we note that the estimation error from the marginal models is taken into account. The seasonal FRG copula was fitted to a suitable rotation of the data, cf. Section 3.2.

Since the FRG copula has three parameters, there are of course different ways of incorporating yearly seasonality in the model. Some alternatives in terms of the FRG copula are discussed in Appendix B, where we also provide detailed estimation results to support the model specification stated above. On a different note, we stress that selecting the static FRG copula as the optimal model amongst static alternatives does not guarantee that the seasonal FRG copula will be preferred to time-varying extensions of other copula models. As a result, similar extensions as those proposed in Eqs. (A.10)–(A.11)

were implemented for most of the copulas in Table A.5 to ensure that the seasonal FRG is superior in terms of AIC.

To illustrate the fit of the proposed seasonal FRG model, we plot in Fig. A.8(a) the empirical Spearman's ρ together with $\hat{\rho}_t^{SFRG}$ implied by the seasonal FRG, using a 60-days moving window, as in Fig. A.6. As a standard of comparison, we include the Spearman's ρ implied by the static FRG. The results indicate that the dependence implied by the seasonal FRG follows the yearly cycle observed in the actual Spearman's ρ rather well. Moreover, it appears from Fig. A.8(a) that we would underestimate the strength of the dependence between the CSS and the LWI during autumn and winter with the static FRG. The reverse is observed during spring and summer, with the strength of the dependence being overestimated by the static FRG. To provide further support for the proposed seasonal FRG, we display in Fig. A.8(b) a simulated path over a four-year horizon, which resembles the actual data nicely.



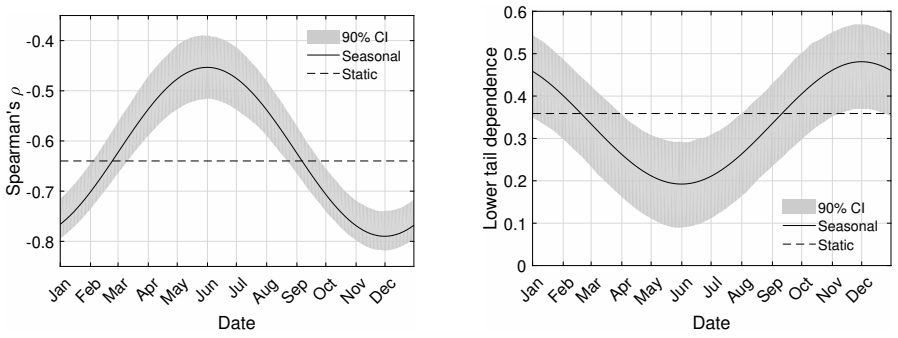
(a) Actual and implied Spearman's ρ

(b) One simulated path from the seasonal FRG

Fig. A.8: (left) Actual Spearman's ρ and Spearman's ρ implied by the static and seasonal FRG copula, based on a 60-days rolling window. The date corresponding to each estimate refers to the last day in the 60-days period. (right) A simulated path of Spearman's ρ from the seasonal FRG copula, aggregated using a 60-days rolling window.

To complement Fig. A.8(a), a clearer picture of the yearly shape of Spearman's ρ implied by the seasonal FRG is given in Fig. A.9(a), where we illustrate the fit at each time point during a year (i.e., no averaging of Spearman's ρ is performed). Equally interesting to consider is the lower tail dependence implied by the fitted seasonal FRG copula, which follows directly from applying Eq. (A.7). The results are plotted in Fig. A.9(b), revealing that the lower tail dependence coefficient reaches its lowest value of approximately 0.2 during summer and its maximum value of approximately 0.50 during winter. This entails that there is a rather large difference between the probability of observing the event of extreme high wind index / low CSS during

winter compared to summer. In Fig. A.9(b), we again provide as benchmark the corresponding static estimate.



(a) Implied Spearman's ρ between CSS and LWI

(b) Implied lower tail dependence between CSS and LWI

Fig. A.9: Spearman's ρ and lower tail dependence implied by the static and seasonal FRG copulas throughout the year. A 90% confidence interval is provided for the seasonal parts.

Dynamic Copula Models

Before proceeding to quantifying the hedging benefits of wind power futures, we comment briefly on another class of models that have become popular in the related literature because of their broad applicability: The Generalized Autoregressive Score (GAS) models proposed by Creal *et al.* (2013). Extending the seasonal copula mixture proposed in Eqs. (A.10)–(A.11) as to allow for one or more parameters to evolve according to the GAS equation is possible. In the context of our study however, where we shall base hedging decisions on simulations over a long time horizon, we argue against such extensions. With the GAS model, we would introduce much complexity compared to the present straightforwardness of the seasonal copula, but not add that much value. Moreover, we stress that compared to the static dependence model, the addition of the simple seasonal extension not only provides a significant improvement, but is very easily interpretable, making it very appealing from a practical perspective.

For applications such as forecasting or short-term simulation, we acknowledge the added value of including a GAS dynamic to the seasonal copula mixture. Therefore, we include in Appendix B a description of the GAS model together with estimation results obtained by fitting two dynamic copula models to our data.

5 Application Results

Having established a model for the joint behavior of the CSS and the wind index, we consider next the quantification of the benefits that WPF can offer GFPPs. A GFPP acting in the day-ahead market can decide from day to day whether to run or not, and thereby take advantage of the daily variation in the CSS. By the construction of the day-ahead wholesale electricity market, a GFPP will not run in times of a negative CSS; its profit R^{CSS} for a period $t \in [T_1, T_2]$ can thus be represented as

$$R^{CSS} = \sum_{t=T_1}^{T_2} 12 \max(CSS_t, 0)s, \tag{A.12}$$

where s is the size of the GFPP measured in MW, and 12 is the number of peak load hours during a day. Recalling the payoff in Eq. (A.2), taking a position $\gamma \in \mathbb{Z}$ in WPF contracts yields the hedged profit of the GFPP, which we denote by R:

$$R = R^{CSS} + \gamma R^{WPF}. \tag{A.13}$$

We note that by excluding weekends and holidays from our analysis, these are not captured in R^{WPF}. We argue however that this does not alter the overall conclusions drawn below. To facilitate hedging decisions, we perform Monte Carlo simulations from the proposed model. Specifically, the marginal models fitted in Section 3.1 and the seasonal FRG copula fitted in Section 4.1 are employed to produce simulations of the joint behavior of the CSS and the wind index, i.e., the pair (CSS_t, W_t). The "price" W_{t_0} affecting R^{WPF} in Eq. (A.13) is computed by averaging across all Monte Carlo simulations of W_t for the delivery $[T_1, T_2]$.[3] While we recognize that this pricing approach is simplistic in that it assumes a zero market price of risk, it simplifies our hedging exercise somewhat, since the mean of the hedged profit R will not be affected by varying the quantity γ. Consequently, instead of the classical mean-variance objective, we can restrict ourselves to the variance minimization criterion in order to determine optimal positions in WPF contracts.

[3]When constructing the wind index data used in this paper, we considered the historical evolution of its two underlying data components, namely the wind power production and the installed capacity, as discussed in Section 2. This implicitly means that we have captured 1) the variations due to changes in wind speeds and 2) the variations caused by the increase in installed capacity and changes in the geographical distribution of wind turbines. While the latter aspect is important to capture in the modeling part of this paper, we argue that a different wind index series should be used in a pricing context. This is because today's WPF price is not affected by the historical evolution of the installed capacity and the changes in the geographical distribution of turbines, but by the *present* conditions. We argue that this issue does not affect the conclusions drawn in this paper, but can have serious implications in other contexts. For more details, we refer the interested reader to Benth and Pircalabu (2018).

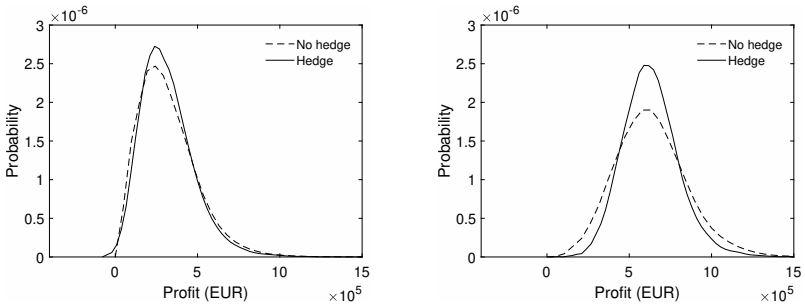
Hence, we consider the following objective:

$$\min_{\gamma \in \mathbb{Z}} \mathbb{V}\text{ar}[R]. \tag{A.14}$$

5.1 Effectiveness of Wind Power Futures

To illustrate the results obtained by applying the hedging approach described above, we fix $s = 200$ MW, $t_0 = 30$ December 2016 (the last date in our sample), and perform 20,000 Monte Carlo simulations of the pair (CSS_t, W_t) one year ahead. The resulting simulated paths are split into monthly periods, and WPF prices corresponding to monthly deliveries are computed as explained earlier. Then, monthly quantities for R^{CSS} and R^{WPF} are constructed for each simulated path, and the minimization in Eq. (A.14) is applied to each month in turn. The subdivision to monthly profits is motivated by the seasonal pattern observed in the dependence structure cf. Fig. A.8(a), and allows us to investigate the effect of the yearly seasonality on hedging-related aspects.

In Fig. A.10, we illustrate the simulated unhedged profit distribution R^{CSS} and the hedged profit distribution R obtained by solving Eq. (A.14) for the months July and October. We observe a compression of the profit distribution in both cases when applying the hedge, which entails that WPF have variance reducing effects. In fact, this finding applies to all 12 months, as will be illustrated shortly.



(a) July (b) October

Fig. A.10: Examples of profit distributions before and after hedging with WPF, based on Monte Carlo simulations from the proposed seasonal FRG copula model.

Perhaps unsettling is the fact that losses can occur when considering the hedged profit distribution in Fig. A.10(a), whereas the unhedged profits cannot attain negative values by construction (see Eq. (A.12)). Nevertheless, we find that the probability of a loss when hedging with WPF is approximately 0.4% on average. The downside of performing the hedge is therefore quite small. In the pursuit to impair this concern even further, recall that the price

of WPF is computed under the assumption of a zero market price of risk. In reality, the studies of Gersema and Wozabal (2017) and Benth and Pircalabu (2018) find evidence of a negative market price of risk in the German market for WPF, implying that a GFPP buys WPF at a discounted price compared to the one computed here. Accounting for this would shift the hedged profit distributions to the right, potentially excluding losses altogether.

Next, we consider in more detail the reduction in the variance of profit distributions attained by performing the hedge. The results are stated in Fig. A.11 for all months of the year, and reveal considerable reductions; even for May and June, where we observe the lowest values, the variance reductions are above 10 %. Further, notice the connection between the yearly pattern of the reductions in Fig. A.11 and the implied Spearman's ρ in Fig. A.9(a): Not surprisingly, the stronger the dependence, the higher the variance reduction.

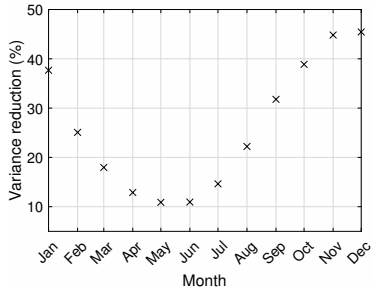


Fig. A.11: Variance reduction achieved by hedging with WPF, for each month of 2017. The results are based on Monte Carlo simulations from the proposed seasonal FRG copula model.

Also relevant to consider in this context is the impact on hedging effectiveness from changing the GFPP efficiency. To assess this, we allow the efficiency to vary from 43.13% to 55.13% with a step size of 2%, and let the emission factor vary according to

$$e = 0.184h,$$

which is based on ICIS (2016). Fig. A.12 illustrates the variance reductions obtained with the different efficiencies, across all months of the year. It appears that increasing the efficiency (i.e., lowering the heat rate) leads to an increase in the variance reductions for all months. The effect seems to be more pronounced during autumn and winter compared to spring and summer. From Eq. (A.1), it was already apparent that increasing the efficiency of a GFPP produces a higher CSS and hence increases profitability. The findings presented in Fig. A.12 incentivize such action even further: Aside from the higher CSS, an increased hedging effectiveness of WPF can be achieved.

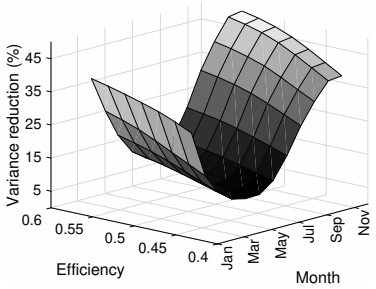


Fig. A.12: Variance reduction implied by hedging with WPF, for GFPP efficiencies spanning from 43.13% to 55.13% with a step size of 2%.

5.2 Comparison with Alternative Models

So far in the hedging application, we have focused on the results obtained with the preferred copula model, that is, the seasonal FRG. In this section, we wish to highlight the benefits of this copula compared to other less optimal alternatives. We consider the following natural progression in comparisons:

1. Frank copula versus FRG copula: In this comparison, we focus on the effect of asymmetry and tail dependence on the benefits of WPF. These features are captured by the FRG copula, as discussed in Section 3.2, but not by the Frank copula, which imposes symmetry and no tail dependence.

2. FRG copula versus seasonal FRG copula: Here, we concentrate on the effect of seasonal dependence on the hedging benefits.

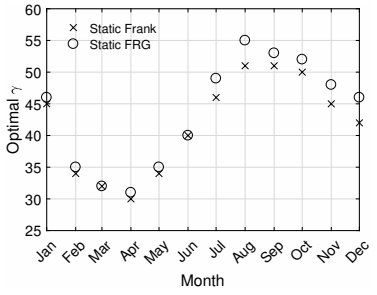
To perform comparisons, we keep the marginal models proposed in Section 3.1 fixed, and repeat the simulations performed in Section 5.1 with the Frank, FRG, and seasonal FRG copulas, instead of only the seasonal FRG. We note that the same random seed was used to produce Monte Carlo simulations from the three models. Then, we compute optimal hedge quantities γ and associated variance reductions with each model, on a monthly basis.

The Effects of Asymmetry and Tail Dependence in the Copula

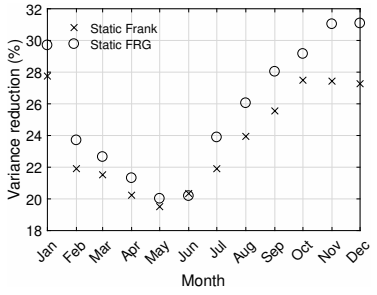
Recalling Figs. A.5(a) and A.5(c), it is apparent that by employing the FRG copula instead of the Frank copula, we introduce a slight asymmetry and assign more probability to the extreme events where high wind and low CSS happen simultaneously. The resulting effects on hedging are depicted in Fig. A.13, where we present the optimal hedge quantities and variance reductions produced by the two copulas. Regarding the former, we notice that the Frank copula generally suggests less WPF in the hedging portfo-

lio. Further, the optimal hedge quantities vary across the year, which is a consequence of the seasonality captured in the marginal models.

Turning to the variance reductions, which are depicted in Fig. A.13(b), we observe that the values implied by the Frank copula are generally lower compared to those implied by the FRG copula. This finding is expected, since GFPPs seek to cover their exposure to high wind / low CSS scenarios. By assigning more probability mass to precisely these events happening simultaneously, which is done by shifting from the Frank to the FRG copula, we increase the benefits of WPF. At the same time, due to the asymmetric behavior of the FRG copula, we are not increasing the probability of observing the reverse combination of low wind / high CSS, and thus not counteracting the increased benefits of WPF. Briefly put, by believing in a dependence structure described by the Frank copula compared to the FRG copula, we would underestimate the risk-reducing power of WPF.



(a) Optimal hedge quantity for each month of 2017 implied by the Frank and FRG copulas



(b) Reduction in the variance of the profit distribution for each month of 2017, when applying the hedges implied by the Frank and FRG copulas

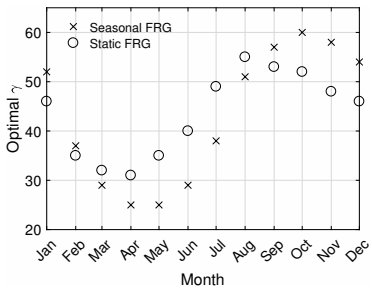
Fig. A.13: Comparison of hedging results implied by the Frank and FRG copulas.

The Effects of Seasonal Time Variation in the Copula

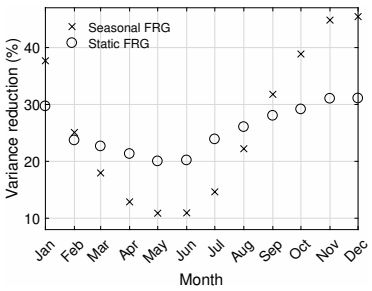
Proceeding to the comparison of the FRG copula with its seasonal version, we present in Fig. A.14 results that are similar to those in Fig. A.13. Regarding the optimal hedge quantities in Fig. A.14(a), the FRG copula yields higher values than the seasonal FRG copula during spring and summer, while the situation reverses during autumn and winter. This alternating behavior is connected to that of the differences in Spearman's ρ implied by the FRG and the seasonal FRG, cf. Fig. A.9(a). That is, the hedge quantities decrease (increase) with a decrease (increase) in absolute values of Spearman's ρ.

Considering Fig. A.14(b), the results reveal a fairly constant level in the variance reduction produced by the FRG copula, compared to the variance

reduction levels implied by the seasonal model. Hence, believing in static dependence can lead to very misleading conclusions when managing risks. Again in this context, we mention the link between the difference in percentage reductions and the difference between the Spearman's ρ implied by the two models (cf. Fig. A.9(a)): The difference in reductions is largest in May/June and November/December, reflecting the fact that the dependence implied by the seasonal FRG model is weakest during May/June and strongest during November/December.



(a) Optimal hedge quantity for each month of 2017 implied by the FRG and seasonal FRG copulas



(b) Reduction in the variance of the profit distribution for each month of 2017, when applying the hedges implied by the FRG and seasonal FRG copulas

Fig. A.14: Comparison of hedging results implied by the FRG and seasonal FRG copulas.

Having found clear evidence of seasonal dependence between the CSS and the wind index, we conclude this section by briefly addressing the error we would get by applying a hedge based on the static FRG model in a seasonal time-varying reality. To perform this analysis, we asses the optimal hedge quantities implied by the static FRG copula in a seasonal setting by using the simulated CSS and wind index from the seasonal model. The variance reductions obtained with this approach are then compared with the reductions implied by the seasonal model shown in Fig. A.14(b). The results, presented in Fig. A.15, reveal very small errors. The smallest and largest errors occur in February and June, respectively, which is connected to the findings presented in Fig. A.14(a); the absolute difference in the optimal number of WPF in the static and seasonal case generates the pattern seen in Fig. A.15.

With the small errors in mind, the real error one commits by believing in static dependence, is the belief in a wrong resulting variance reduction. Thus, while the static model creates a misleading picture in a risk management context, our results suggest that it could be employed to determine optimal hedging quantities.

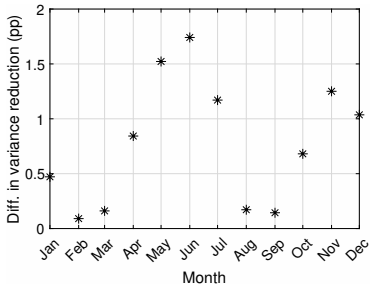


Fig. A.15: The effects of using the "wrong" hedge quantity: Difference in variance reduction of the profit distribution when using hedge quantities obtained from 1) the FRG, and 2) the seasonal FRG, both evaluated using simulations from the seasonal FRG copula model.

5.3 Discussion of the Proposed Hedging Strategy

To conclude this section, we turn briefly to the standard hedging principle often employed to hedge the day-ahead CSS. Usually, conventional generators remove their exposure to day-ahead price risk either completely or partially by entering a short position in standard power forwards, and a long position in fuel forwards and carbon credit forwards. In this paper, we have considered a different hedging approach with the purpose of determining the potential of the newly introduced wind power futures, but we stress that our strategy is not incompatible with the industry standard. In fact, our hedging portfolio consisting of wind power futures could be extended to include the additional forwards mentioned above. However, this would require us to switch from our bivariate modeling problem to a multivariate one, since the joint behavior of the wind index, the day-ahead CSS, and the different forward clean spark spreads should be considered. While this is outside the scope of the present paper, it is nevertheless an interesting perspective that has not been studied yet, and could possibly be approached with vine copulas. In the context of the effectiveness of the standard hedging principle that conventional generators usually employ, we mention the study of Charalampous and Madlener (2016).

6 Conclusion

In this paper, we propose a joint model for the day-ahead clean spark spread and the daily wind index that can facilitate hedging decisions for a gas-fired power plant. The modeling procedure is based on two steps: First, the marginal behavior of the variables is considered, where we apply seasonal functions and ARMA–GARCH filters to remove the seasonality and the serial dependence in the conditional mean and variance. While the usual Gaussian

assumption for the innovation process works in the case of the daily wind index, the normal-inverse Gaussian distribution provides a better fit for the day-ahead clean spark spread. Second, the standardized residuals from the ARMA–GARCH models are connected through copulas. The data reveals a dependence structure that is slightly asymmetric, and also varying according to an annual cycle. To capture these empirical findings, we propose a seasonal copula mixture, where the mixing components are the rotated Gumbel and the Frank copulas.

Based on Monte Carlo simulations from the proposed model, we show that wind power futures have considerable risk-reducing benefits in the context of a gas-fired power plant operating in the day-ahead market. Further, their hedging effectiveness increases as a function of the efficiency of the gas-fired power plant. To highlight the importance of capturing asymmetry, tail dependence, and seasonality in the dependence structure, we perform comparison studies where the optimal model is compared to less optimal alternatives. Accounting for asymmetry and tail dependence (as opposed to imposing symmetry and zero tail dependence) leads to an increase in the effectiveness of wind power futures. Moreover, we find that the conclusions drawn with a static dependence model deviate to a large extent from those obtained with a seasonal dependence model. With static dependence, the variance reductions of the profit distributions attained by the hedge vary between 20% and 31%; in the seasonal case the corresponding reductions vary between 10% and 45%.

Although we have concentrated on the German market and the case of gas-fired power plants, the results are relevant for other markets, and are also transferable to other conventional electricity producers. Since the amount of electricity generated by wind turbines is expected to grow globally, the dependence between the day-ahead clean spark spread and the daily wind index in other market places will most likely be strengthened in the future. Hence, it is reasonable to assume that more weather-based instruments similar to the German wind power futures will be introduced, enabling similar analyses to be performed on other than the German market.

Acknowledgements

The authors would like to thank the Quantitative Analytics team at Neas Energy and the participants attending the Second Conference on the Mathematics of Energy Markets 2017 at the Wolfgang Pauli Institute for providing valuable comments and suggestions. Three anonymous referees are also thanked for their constructive criticism and suggestions, which improved the presentation of this paper.

Funding

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B. Anca Pircalabu is supported by the Innovation Fund Denmark under Grant 4135-00082B.

A Theoretical Results for Copula Mixtures

A.1 Proof of Proposition 1

For positive dependent variables, the upper tail dependence for copula C can be written as (see e.g. (McNeil *et al.*, 2005, p. 209))

$$\Lambda^u = \lim_{u \uparrow 1} \frac{\hat{C}(1-u, 1-u)}{1-u} = \lim_{u \uparrow 1} \frac{1 - 2u + C(u,u)}{1-u}, \tag{A.15}$$

where $\hat{C}(u,v) := P(U > u, V > v) = C(1-u, 1-v) + u + v - 1$ is the survival copula. Applying Eq. (A.15) to the copula mixture, we get that the upper tail dependence is

$$\begin{aligned}
\Lambda^u =& \lim_{u \uparrow 1} \left[\frac{1 - 2u + \lambda C^a(u,u) + (1-\lambda) C^b(u,u)}{1-u} \right] \\
=& \lim_{u \uparrow 1} \left[\frac{1-2u}{1-u} + \lambda \frac{1 - 2u + C^a(u,u)}{1-u} + (1-\lambda) \frac{1 - 2u + C^b(u,u)}{1-u} \right. \\
& \left. - \lambda \frac{1-2u}{1-u} - (1-\lambda) \frac{1-2u}{1-u} \right] \\
=& \lambda \Lambda^{u,a} + (1-\lambda) \Lambda^{u,b}.
\end{aligned}$$

Similarly, the lower tail dependence can be written in terms of copula C as (again for positive dependent variables)

$$\Lambda^l = \lim_{u \downarrow 0} \frac{C(u,u)}{u},$$

resulting in the following lower tail dependence for the copula mixture:

$$\begin{aligned}
\Lambda^l &= \lim_{u \downarrow 0} \left[\frac{\lambda C^a(u,u) + (1-\lambda) C^b(u,u)}{u} \right] \\
&= \lambda \Lambda^{l,a} + (1-\lambda) \Lambda^{l,b}.
\end{aligned}$$

A.2 Proof of Proposition 2

From Eq. (A.8) we have that Spearman's ρ implied by copula C^m is

$$\begin{aligned}
\rho^m =& 12 \int_0^1 \int_0^1 C^m(u_1, u_2) \mathrm{d}u_1 \mathrm{d}u_2 - 3 \\
=& 12 \int_0^1 \int_0^1 \left(\lambda C^a(u_1, u_2) + (1-\lambda) C^b(u_1, u_2) \right) \mathrm{d}u_1 \mathrm{d}u_2 - 3(1 - \lambda + \lambda) \\
=& \lambda \left(12 \int_0^1 \int_0^1 C^a(u_1, u_2) du_1 du_2 - 3 \right) \\
& + (1-\lambda) \left(12 \int_0^1 \int_0^1 C^b(u_1, u_2) du_1 du_2 - 3 \right) \\
=& \lambda \rho^a + (1-\lambda) \rho^b.
\end{aligned}$$

B Alternative Models

As mentioned in Section 4.1, there are many ways of extending the FRG copula. Here, we present some alternatives to the model proposed in Eqs. (A.10)–(A.11), and the corresponding estimation results.

B.1 Seasonal Copula Alternatives

Instead of introducing a yearly cycle in the evolution equation for Spearman's ρ corresponding to the RGumbel copula cf. Eq. (A.10), one could consider this for Frank, i.e.,

$$\rho_t^F = a^F + b^F \sin(2\pi t/K) + c^F \cos(2\pi t/K). \tag{A.16}$$

Alternatively, yearly seasonality could be introduced in both Spearman's ρ implied by the Frank and the RGumbel copula. These alternative models are stated in Table A.7, and corresponding estimation results are reported in Table A.8.

The results in Table A.8 reveal that the seasonal FRG model proposed in Section 4.1 is slightly better than Alternative 1 in terms of AIC. Considering Alternative 2, although its AIC is slightly lower than that of the other seasonal models, notice the large standard errors, which imply that the seasonal parameters corresponding to the Frank part are not statistically significant at a 5% level.

B.2 Dynamic Copula Alternatives

In terms of dynamic copula alternatives, we consider the Generalized Autoregressive Score (GAS) model proposed by Creal *et al.* (2013). Assuming a

Seasonal FRG	Alternative 1
$\rho_t^{SFRG} = \lambda^{SFRG}\rho^F + (1-\lambda^{SFRG})\rho_t^{RG}$	$\rho_t^{SFRG} = \lambda^{SFRG}\rho_t^F + (1-\lambda^{SFRG})\rho^{RG}$
$\rho_t^{RG} = a^{RG} + b^{RG}\sin\left(\frac{2\pi t}{K}\right) + c^{RG}\cos\left(\frac{2\pi t}{K}\right)$	$\rho_t^F = a^F + b^F\sin\left(\frac{2\pi t}{K}\right) + c^F\cos\left(\frac{2\pi t}{K}\right)$
Alternative 2	
$\rho_t^{SFRG} = \lambda^{SFRG}\rho_t^F + (1-\lambda^{SFRG})\rho_t^{RG}$	
$\rho_t^{RG} = a^{RG} + b^{RG}\sin\left(\frac{2\pi t}{K}\right) + c^{RG}\cos\left(\frac{2\pi t}{K}\right)$	
$\rho_t^F = a^F + b^F\sin\left(\frac{2\pi t}{K}\right) + c^F\cos\left(\frac{2\pi t}{K}\right)$	

Table A.7: Model specifications for the seasonal FRG proposed in Section 4.1 and two seasonal alternatives.

	$\hat{a}^F$	$\hat{b}^F$	$\hat{c}^F$	$\hat{a}^{RG}$	$\hat{b}^{RG}$	$\hat{c}^{RG}$	λ^{SFRG}	$\log \mathcal{L}_c$	AIC
Seasonal FRG	0.742	-	-	0.566	-0.132	0.208	0.316	309.894	-609.788
	(0.072)	(-)	(-)	(0.041)	(0.040)	(0.043)	(0.092)		
Alternative 1	0.573	-0.122	0.280	0.689	-	-	0.532	309.657	-609.314
	(0.048)	(0.048)	(0.056)	(0.054)	(-)	(-)	(0.097)		
Alternative 2	0.693	0.050	0.156	0.580	-0.162	0.150	0.374	312.454	-610.908
	(0.080)	(0.120)	(0.099)	(0.055)	(0.081)	(0.071)	(0.109)		

Table A.8: Maximum likelihood estimation results for the models stated in Table A.7. The maximized value of the copula log-likelihood is denoted $\log \mathcal{L}_c$. Standard errors are reported in parenthesis and are computed following the simulation-based procedure in Patton (2013).

copula having one governing parameter which we denote by θ, the evolution equation for the GAS(1,1) is given by

$$z_{t+1} = \omega + \Psi z_t + \varsigma I_t^{-1/2} s_t, \tag{A.17}$$

where

$$\begin{aligned} z_t &= f(\theta_t) \\ s_t &= \frac{\partial}{\partial \theta_t} \log c(u_{1,t}, u_{2,t} | \theta_t) \\ I_t &= \mathbb{E}[s_t^2], \end{aligned}$$

and ω, Ψ, and ς are constant coefficients.

Recalling Eq. (A.5) and restricting this analysis to the FRG copula, we let the (transformed) parameter of the RGumbel copula evolve according to the GAS(1,1) model. Specifically,

$$\begin{aligned} c_t^{FRG}(\cdot|\mathcal{F}_{t-1}, \theta^F, \theta_t^{RG}, \lambda^{FRG}) =& \lambda^{FRG} c^F(\cdot|\mathcal{F}_{t-1}, \theta^F) \\ &+ (1 - \lambda^{FRG}) c_t^{RG}(\cdot|\mathcal{F}_{t-1}, \theta_t^{RG}), \end{aligned}$$

where $z_t = \log(\theta_t^{RG} - 1)$ follows a GAS(1,1). Aside from this, we consider two other alternatives which are stated in Table A.9; one where we only allow the transformed RGumbel copula parameter to vary according to a yearly cycle, and one where we combine the GAS dynamics with the yearly seasonality.

Alternative 3 (GAS)	Alternative 4 (Season)
$z_{t+1} = \omega^{RG} + \Psi^{RG} z_t + \varsigma^{RG} I_t^{-1/2} s_t$	$z_{t+1} = \omega^{RG} + b^{RG} \sin\left(\frac{2\pi t}{K}\right) + c^{RG} \cos\left(\frac{2\pi t}{K}\right)$
Alternative 5 (Season GAS)	
$z_{t+1} = \omega^{RG} + \Psi^{RG} z_t + \varsigma^{RG} I_t^{-1/2} s_t + b^{RG} \sin\left(\frac{2\pi t}{K}\right) + c^{RG} \cos\left(\frac{2\pi t}{K}\right)$	

Table A.9: Model specifications for three alternative models.

Parameter estimates for the models stated in Table A.9 are obtained by maximum likelihood, and are reported in Table A.10. Comparing Alternative 3 to Alternative 4, AIC decreases quite a bit when considering a deterministic yearly cycle instead of the GAS(1,1) specification. Further, combining the two effects (Alternative 5) yields the lowest AIC value across all models; notice however that two out of seven parameters are not significant at a 5% level.

Regarding Alternative 4, we stress that this model is very similar to the seasonal FRG: in Alternative 4, we let the transformed copula parameter vary according to a yearly cycle, whereas in the seasonal FRG, we introduce yearly seasonality in Spearman's ρ. Although the AIC corresponding to the former specification is slightly lower, we find the seasonal FRG to be more appealing, given that the yearly cycle is detected in the rolling Spearman's ρ cf. Fig. A.6.

	$\hat{\theta}^F$	$\hat{\omega}^{RG}$	$\hat{\Psi}^{RG}$	$\hat{\varsigma}^{RG}$	$\hat{b}^{RG}$	$\hat{c}^{RG}$	$\hat{\lambda}^{FRG}$	$\log \mathcal{L}_c$	AIC
Alternative 3 (GAS)	6.132	0.001	0.993	0.092	-	-	0.315	300.349	-590.697
	(2.393)	(0.648)	(0.378)	(0.314)	(-)	(-)	(0.262)		
Alternative 4 (Season)	6.605	-0.391	-	-	-0.451	0.758	0.312	310.376	-610.752
	(1.776)	(0.148)	(-)	(-)	(0.144)	(0.152)	(0.090)		
Alternative 5 (Season GAS)	6.725	-0.421	-0.231	0.235	-0.503	0.842	0.270	315.832	-617.663
	(2.831)	(0.379)	(0.567)	(0.107)	(0.208)	(0.303)	(0.121)		

Table A.10: Maximum likelihood estimation results for the models stated in Table A.9. The maximized value of the copula log-likelihood is denoted $\log \mathcal{L}_c$. Standard errors are reported in parenthesis and are computed following the simulation-based procedure in Patton (2013).

References

Aepli, M. D., Füss, R., Henriksen, T. E. S. and Paraschiv, F. (2017). Modeling the multivariate dynamic dependence structure of commodity futures portfolios. *Journal of Commodity Markets*, **6**, 66–87.

Avdulaj, K. and Barunikl, J. (2015). Are benefits from oil - stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, **51**, 31–44.

Barndorff-Nielsen, O. E. (1997a). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, **24**(1), 1–13.

Barndorff-Nielsen, O. E. (1997b). Processes of normal inverse Gaussian type. *Finance and Stochastics*, **2**(1), 41–68.

Benth, F. E. and Šaltytė Benth, J. (2004). The normal inverse Gaussian distribution and spot price modelling in energy markets. *International Journal of Theoretical and Applied Finance*, **07**(02), 177.

Benth, F. E. and Šaltytė Benth, J. (2011). Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, Vol. 2011, Article ID 576791, 21 pages.

Benth, F.E. and Kettler, P. C. (2011). Dynamic copula models for the spark spread. *Quantitative Finance*, **3**, 407–421.

Benth, F. E. and Pircalabu, A. (2018). A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures. *Applied Mathematical Finance*. **25**(1), 36–65.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *European Journal of Finance*, **15**(7–8), 675–701.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**(3), 307–327.

Börger, R., Cartea, A., Kiesel, R. and Schindlmayr, G. (2009). Cross-commodity analysis and applications to risk management. *Journal of Futures Markets*, **29**(3), 197–217.

Charalampous, G. and Madlener, R. (2016). Risk management and portfolio optimization for gas- and goal-fired power plants in Germany: A multivariate GARCH approach. *Journal of Energy Markets*, **9**(2), 69–94.

Creal, D., Koopman, S.J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, **28**, 777-795.

Dias, A. and Embrechts, P. (2009). Testing for structural changes in exchange rates' dependence beyond linear correlation. *The European Journal of Finance*, **15**(7–8), 619–637.

Elberg, C. and Hagspiel, S. (2015). Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, **241**, 260–272.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**(4), 987–1007.

European Commission, The EU Emissions Trading System, 2017. `https://ec.europa.eu/clima/policies/ets_en` (accessed 29 August 2017).

Genest, C., Remillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, **44**(2), 199–213.

Gersema, G. and Wozabal, D. (2017). An equilibrium pricing model for wind power futures. *Energy Economics*, **65**, 64–74.

Grothe, O. and Schneiders, J. (2011). Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, **39**, 4742–4754.

Haldrup, N. and Nielsen, M. Ø., 2006. A regime switching long memory model for electricity prices. *Journal of Econometrics*, **135**(1–2), 349–376.

Härdle, W.K. and López Cabrera, B. (2012). Implied market price of weather risk. *Applied Mathematical Finance*, **19**(1), 59–95.

ICIS (2016). European Daily Electricity Markets Methodology.

Jensen, M. B. and Lunde, A. (2001). The NIG-S&ARCH model: A fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model. *The Econometrics Journal*, **4**(2), 319–342.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, **94**(2), 401–419.

McNeil, A. J., Frey, R. and Embrechts, P. (2005). Quantitative risk management. *Princeton Series in Finance*.

Nelsen, R.B. (2006). *An Introduction to Copulas*, Springer.

Patton, A.J. (2006(a)). Modelling asymmetric exchange rate dependence. *International Economic Review*, **47**(2), 527–556.

Patton, A.J. (2006(b)). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, **21**(2), 147–173.

Patton, A.J. (2013). Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899 – 960. Elsevier B.V.

Pircalabu, A. and Jung, J. (2017). A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quantitative Finance*, **17**(10), 1583–1600.

Pircalabu, A., Hvolby, T., Jung, J. and Høg, E. (2017). Joint price and volumetric risk in wind power trading: A copula approach. *Energy Economics*, **62**, 139–154.

Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance*, **14**(3), 401–423.

Rydberg, T. H. (1997). The normal inverse Gaussian Lévy process: Simulation and approximation. *Communications in Statistics. Stochastic Models*, **13**(4), 887–910.

Shumway, R. H. and Stoffer, D. S. (2011). Time Series Analysis and Its Applications. *Springer Texts in Statistics*.

Sklar, A. (1059). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, **8**, 229–231.

Paper B

On the Spatial Hedging Effectiveness of German Wind Power Futures for Wind Power Generators

List of authors: Troels Sønderby Christensen[1,2], Anca Pircalabu[1,2]

[1]Department of Mathematical Sciences, Aalborg University
[2]Quantitative Analytics, Centrica Energy Trading

The paper has been published in
Journal of Energy Markets Vol. 11(3), pp. 71–96, September 2018.

The layout has been revised.

Abstract

The wind power futures recently introduced on the German market fill the gap of a standardized product that addresses directly the volume risk in wind power trading. While the German wind power futures entail risk-reducing benefits for wind power generators generally speaking, it remains unclear the extent of these benefits across wind farms with different geographical locations. In this paper, we consider the wind utilization at 31 different locations in Germany, and for each site, we propose a copula model for the joint behavior of the site-specific wind index and the overall German wind index. Our results indicate that static mixture copulas are preferred to the stand-alone copula models usually employed in the economic literature. Further, we find evidence of asymmetric dependence and upper tail dependence. To quantify the benefits of wind power futures at each wind site, we perform a minimum variance hedge, and find that variance reductions can differ greatly depending on the geographical location. Further, different comparison studies reveal that the presence of 1) a negative risk premium in the wind power futures market and 2) upper tail dependence weaken the benefits of wind power futures for wind power generators.

1 Introduction

Wind power generators worldwide have historically been given subsidies in order to incentivize the development of renewable energy sources. The non-programmable nature makes investment in wind power generation unpredictable, and although subsidies simplify investment decisions, the stochastic behavior of wind demands further risk reducing opportunities. As a result, the so-called German wind power futures have recently been introduced on the European Energy Exchange and Nasdaq. These instruments are written on a wind power production index that reflects the average German utilization, and a long (short) position gives a profit in high-wind (low-wind) scenarios. It follows naturally that wind power generators constitute the seller group, as they seek to cover their exposure to the low-wind scenarios that affect their cash-flows negatively.

In this paper, we study the hedging benefits of wind power futures for wind turbines or wind farms with different geographical locations in Germany. To facilitate optimal hedging decisions, we employ copula models to analyze the joint behavior of the site-specific and the German wind power production indexes. Specifically, we base our empirical analysis on data from 31 different sites in Germany, and quantify the hedging benefits of wind power futures for each site, showcasing how these vary across locations. Further, we comment on how the negative risk premium in the wind power futures market affects the wind power generator's hedged profit distribution. For the data we consider, we find significant evidence of asymmetric depen-

dence and upper tail dependence between the site-specific and the German wind utilization. To highlight the importance of capturing such effects, we include a comparison study where the hedging exercise is performed with the optimal copula model and a less optimal alternative that imposes symmetry and no tail dependence.

Owing to the recent introduction of wind power futures to the market, the related literature is yet scarce. The first studies to consider the German wind power futures are Gersema and Wozabal (2017) and Benth and Pircalabu (2017). Both papers concentrate on the pricing rather than the hedging aspect, and agree on the presence of a negative risk premium which is explained by the fact that the wind power producer's profit is more correlated to the German wind utilization than the conventional generator's profit. Considering the hedging aspect is Christensen *et al.* (2019), where the benefits of wind power futures are studied in the context of conventional generators. Due to the prioritization of the cheapest energy sources in the day-ahead market, conventional generators are affected by the share of wind power in the system. In high-wind scenarios, conventional generators will produce less, and they can minimize this exposure by taking a long position in wind power futures. While the present paper is similar to the study in Christensen *et al.* (2019) in that it investigates the hedging power of wind power futures, it differentiates itself by considering the seller rather than the buyer side. Furthermore, we consider multiple potential sellers in order to emphasize the spatial aspect, which in contrast to conventional generators is very important to consider here.

Turning to copulas, their application in the energy markets literature has grown tremendously over the past years, see e.g. Benth and Kettler (2011), Grothe and Schneiders (2011), Avdulaj and Barunikl (2015), Elberg and Hagspiel (2015), Pircalabu and Jung (2017), Aepli *et al.* (2017) and Liu *et al.* (2017) to name a few. Closest to the present paper are the studies of Grothe and Schneiders (2011) and Elberg and Hagspiel (2015), who also employ copulas to model the German wind power production. The study of Grothe and Schneiders (2011) concentrates on the optimal geographical allocation of wind turbines, and Elberg and Hagspiel (2015) quantify the value of wind turbines at different locations in Germany, which requires modeling the dependence between the aggregated German wind power generation and the wind power generation from each wind turbine. While our modeling approach is very similar to that in Elberg and Hagspiel (2015), we introduce further flexibility by considering copula mixtures.

The remaining of this paper is structured as follows: Section 2 presents the data and introduces the wind power futures. In Section 3 we briefly describe the modeling approach and present the estimation results. In Section 4 we examine the spatial hedging effectiveness of wind power futures and provide a comparison study. Section 5 concludes.

2 Data Presentation and the Wind Power Futures

The empirical study performed in this paper relies on wind index data from 31 different sites in Germany, and on the overall German wind index, which acts as the underlying of wind power futures. In the following, we elaborate on the data and introduce the German wind power futures.

2.1 German and Site-Specific Wind Index Data

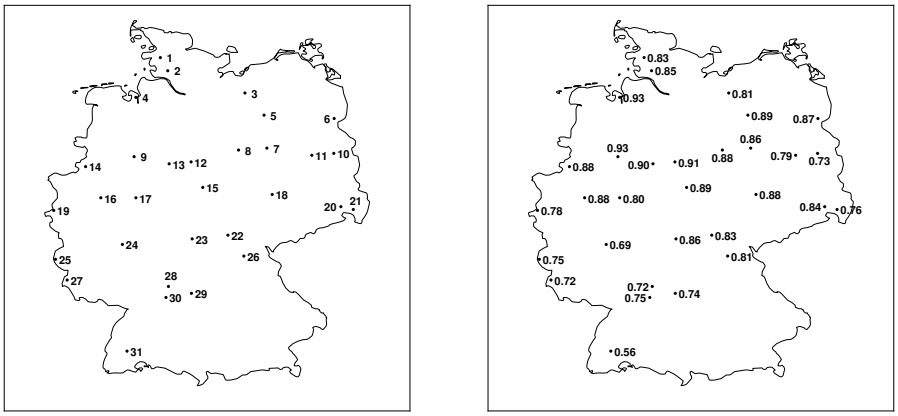
The first data component in our analysis consists of German wind index data provided by Narex. The index represents the aggregated utilization of wind power plants in Germany on a daily basis, and covers the period from 1 January 2012 to 31 December 2015, that is we have a total of 1461 observations. We remark that the index is computed based on wind speed data and a *constant* level of installed capacity in Germany, as to exclude effects that originate from the historical variations in the installed capacity. Specifically, the level is set to that of December 2015, and the German wind index data reflects therefore the distribution of wind generation capacity at that time, since we intend to illustrate the hedging effectiveness of wind power futures contracts in the present conditions.

The second data component consists of wind index data at 31 specific wind sites, again for the period 1 January 2012 to 31 December 2015. For each site, we have access to the actual daily wind power generation data and the installed capacity data. To construct the site-specific wind index, we consider the following: Let $P_{t,i}$ denote the day t wind power generation measured in MWh for wind site i, and let C_i denote the installed capacity for wind site i. The wind index $W_{t,i} \in [0,1]$ is then obtained as

$$W_{t,i} = \frac{P_{t,i}}{h_t C_i}, \tag{B.1}$$

where h_t is the number of hours in a given day. We note that the number of turbines included in each wind site and their type vary across the sites, but not over time. Further, the condition of the wind turbines and the surrounding terrain vary across the sites. While we acknowledge that circumstances other than the geographical location can have an impact on the site-specific wind index, it is not possible to separate such effects based on the data we have available. Therefore, whenever we refer to differences between the wind sites which are caused by their geographical locations, we implicitly mean other circumstances as well.

To illustrate how the 31 wind sites are spread across Germany, we plot in Fig. B.1(a) their approximate geographical location, with corresponding site ID number. Moreover, we report in Fig. B.1(b) the linear correlation between



(a) Location and site ID

(b) Linear correlation

Fig. B.1: Location of wind sites with site ID, and linear correlations between German wind index and each site wind index.

the wind index of each site and the German wind index – not surprisingly, we generally find a very strong positive relation.

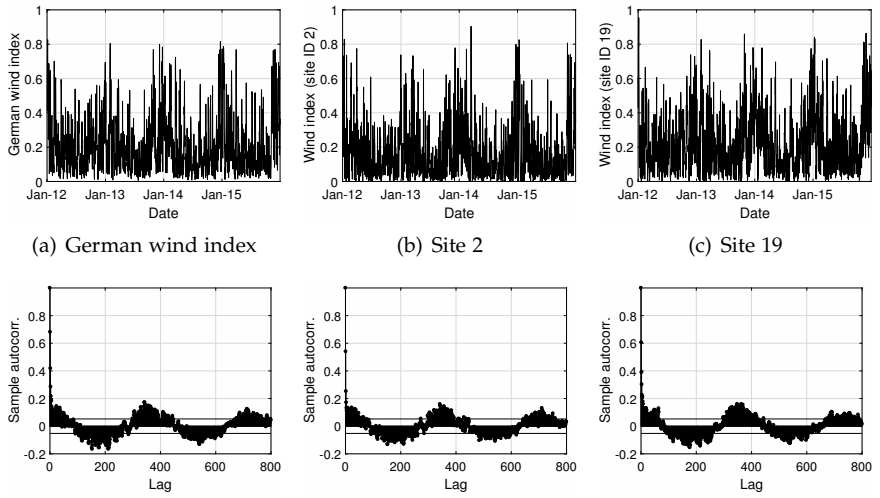
To provide the reader with more sense of the data, we present in the upper row block of Fig. B.2 time series plots of the German wind index and the wind indexes at sites 2 and 19. In the lower row block of Fig. B.2, we plot the corresponding sample autocorrelation functions. The plots reveal clear yearly cycles for the German wind index and the wind indexes at site 2 and 19.

2.2 Wind Power Futures

The German wind index displayed in Fig. B.2(a) acts as the underlying for German wind power futures (WPF) traded on the European Energy Exchange and Nasdaq. To clarify the WPF payoff, let us consider a contract with delivery period $[T_1, T_2]$, where $T_1 < T_2$. Further, let us denote by W_t^{DE} the realized wind index in Germany at day t. Then, a long position in a WPF contract yields the payoff

$$R^{\mathrm{WPF}} = H\left(\frac{1}{T_2 - T_1 + 1}\sum_{t=T_1}^{T_2} W_t^{\mathrm{DE}} - W_{t_0}\right) X, \tag{B.2}$$

where H is the number of wind production hours (wph) in the delivery period, W_{t_0} is the "market price" set at day $t_0 < T_1$, and X is the tick size set to 100 EUR at EEX. The price is quoted in EUR per wph, and a price of e.g. 30 EUR/wph translates to $W_{t_0} = 0.30$.



(a) German wind index (b) Site 2 (c) Site 19

(d) Autocorr. German wind index (e) Autocorr. site 2 (f) Autocorr. site 19

Fig. B.2: Upper row block: Time series plots of the German wind index and the wind index at sites 2 and 19. Lower row block: Sample autocorrelation of the German wind index and the wind index at sites 2 and 19.

It follows from Eq. (B.2) that a long position in a WPF contract generates a profit in high wind scenarios, whereas a short position generates a profit in low wind scenarios. For a wind power generator, a short position in WPF contracts is therefore an interesting opportunity to mitigate the volumetric risk associated with wind power generation.

3 Modeling Procedure

To model the joint behavior of the German wind index and a site-specific wind index, we employ ARMA–GARCH copula models; with 31 different wind sites, we end up with 31 bivariate models. The choice of modeling approach is motivated by the well-known flexibility of these models, and also by the copula being an excellent tool for modeling nonlinear dependence. Further ARMA–GARCH models have proven to perform well for univariate wind-driven stochastic variables such as the wind index, see e.g. Tol (1997) and Šaltytė Benth and Benth (2010). Since ARMA–GARCH copula models are common in the financial and econometric literature, we shall not provide a comprehensive introduction here, and refer the interested reader to e.g. Patton (2006, 2013), Dias and Embrechts (2009), and Pircalabu *et al.* (2017).

In the following subsections, we consider 1) univariate ARMA–GARCH

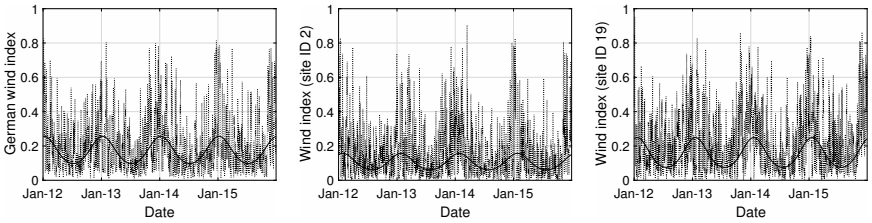
type models for the marginal behavior of the German wind index and all site-specific wind indexes, and 2) bivariate copulas for the dependence between the German wind index and each site-specific wind index. For readability, we only report the results for four chosen sites, namely site 2, 19, 20, and 31 cf. Fig. B.1(a). The four sites are selected as to emphasize the spatial effects.

3.1 Marginal Models

As already illustrated in Fig. B.2, the wind index data is bounded between zero and one. Since ARMA-GARCH models are not suited for modeling this type of bounded data, we first apply the logit transform to the wind indexes.[1] Then, to capture the yearly seasonality revealed in Fig. B.2, we employ a seasonal function of the form,

$$f_t = a_1 + c_1 \sin(2\pi t/365) + c_2 \cos(2\pi t/365), \tag{B.3}$$

which is fitted to each logit-transformed wind index. The resulting seasonal fits are shown in Fig. B.3 for the German wind index, site 2, and site 19.



(a) German wind index (b) Site 2 (c) Site 19

Fig. B.3: Time series plots of the German wind index and the wind index at sites 2 and 19 with fitted seasonal functions.

After removing the yearly seasonality from the data, we filter each marginal series through ARMA–GARCH models. To determine the optimal order of the models, we employ the Bayesian Information Criterion (BIC), where we consider ARMA(p,q)–GARCH(h,k) models of order $p, q = 0, .., 5$ and $(h, k) = \{(0,0), (0,1), (1,1)\}$. Table B.1 presents the optimal models and corresponding parameter estimates with standard errors in parenthesis, revealing that the preferred specifications vary across the chosen sites.

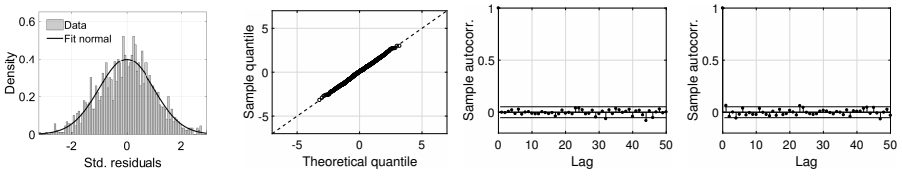
Considering the goodness-of-fit of the proposed models in Table B.1, let us denote by η_t^{DE} the standardized residuals corresponding to the German wind index, and η_t^i the standardized residuals corresponding to the wind index at site i. Diagnostic plots are displayed in Fig. B.4 for the German index, and in Fig. B.5 for four selected sites. The autocorrelation functions

[1] The logit transform is given by $\text{logit}(x) = \log(x) - \log(1-x)$.

	Logit W^{DE}	Logit $W^{\text{Site 2}}$	Logit $W^{\text{Site 19}}$	Logit $W^{\text{Site 20}}$	Logit $W^{\text{Site 31}}$
Model	ARMA(1,1)	ARMA(1,1)–ARCH(1)	AR(1)	AR(2)–ARCH(1)	AR(3)–GARCH(1,1)
			Conditional mean		
AR1 $\hat{\phi}_1$	0.444 (0.039)	0.349 (0.042)	0.519 (0.018)	0.615 (0.036)	0.504 (0.025)
AR2 $\hat{\phi}_2$	-	-	-	-0.086 (0.029)	-0.053 (0.030)
AR3 $\hat{\phi}_3$	-	-	-	-	0.120 (0.026)
MA1 $\hat{\theta}_1$	0.253 (0.041)	0.157 (0.047)	-	-	-
Variance $\hat{\sigma^2}$	0.727 (0.028)	-	1.284 (0.031)	-	-
			Conditional variance		
Constant $\hat{\omega}$	-	1.590 (0.055)	-	1.131 (0.047)	0.046 (0.018)
ARCH $\hat{\alpha}$	-	0.126 (0.029)	-	0.311 (0.031)	0.031 (0.008)
GARCH $\hat{\beta}$	-	-	-	-	0.947 (0.015)
			Marginal distribution		
DoF $\hat{\nu}$	-	6.447 (0.931)	6.734 (1.006)	6.563 (1.063)	8.111 (1.680)
Skewness $\hat{\lambda}$	-	-0.321 (0.032)	-0.215 (0.033)	-0.322 (0.036)	-0.278 (0.033)

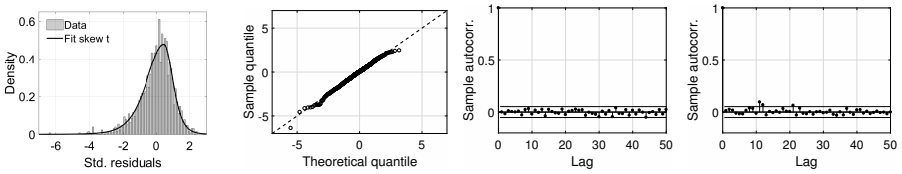
Table B.1: Type and order of marginal models, parameter estimates, and corresponding standard errors in parenthesis.

for the standardized residuals and the standardized residuals squared show almost no correlation, hence indicating appropriate model selection. Regarding the distributional assumptions, the normal distribution provides a nice fit to the standardized residuals corresponding to the German wind index, as it appears from Figs. B.4(a) and B.4(b). For the site-specific wind indexes however, the normal distribution is less appropriate. As a result, we relax the normality assumption and find that a skew t distribution provides a much better fit, as shown in Fig. B.5. Parameter estimates for the fitted skew t distributions together with standard errors are reported in the last row block of Table B.1. Although results concerning four selected sites are displayed in this section, we stress that very similar results are obtained for the remaining series.

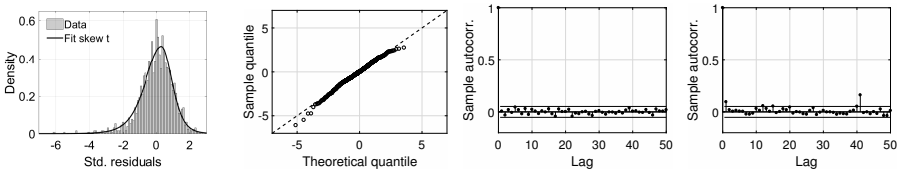


(a) Histogram (b) Quantile plot (c) Sample acf of $\hat{\eta}_t^{DE}$ (d) Sample acf of squared $\hat{\eta}_t^{DE}$

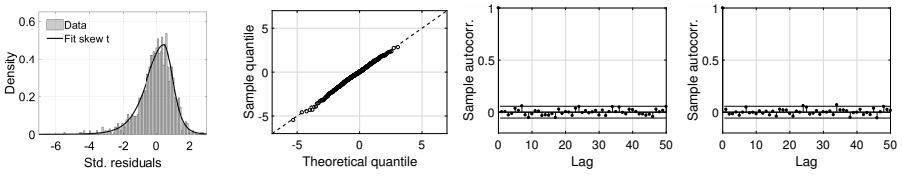
Fig. B.4: Model check for the German wind index.



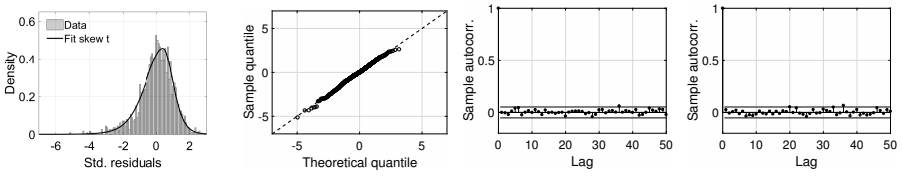
(a) Histogram (site 2) (b) Quantile plot (site 2) (c) Sample acf of $\hat{\eta}_t^i$ (i=2) (d) Sample acf of sq. $\hat{\eta}_t^i$ (i=2)



(e) Histogram (site 19) (f) Quantile plot (site 19) (g) Sample acf of $\hat{\eta}_t^i$ (i=19) (h) Sample acf of sq. $\hat{\eta}_t^i$ (i=19)



(i) Histogram (site 20) (j) Quantile plot (site 20) (k) Sample acf of $\hat{\eta}_t^i$ (i=20) (l) Sample acf of sq. $\hat{\eta}_t^i$ (i=20)



(m) Histogram (site 31) (n) Quantile plot (site 31) (o) Sample acf of $\hat{\eta}_t^i$ (i=31) (p) Sample acf of sq. $\hat{\eta}_t^i$ (i=31)

Fig. B.5: Model check for four selected sites: 2, 19, 20, and 31.

3.2 Copula Models

Next, we turn to dependence modeling with copulas, which are esentially distributions with uniform margins. Let us consider the random vector $\boldsymbol{Y}_t = (Y_{t,1}, Y_{t,2})$, with joint conditional distribution $F(\cdot|\mathcal{F}_{t-1})$. Then, according to Sklar's theorem [Sklar (1959)] for conditional distributions, we can decompose $F(\cdot|\mathcal{F}_{t-1})$ as

$$F(y_{t,1}, y_{t,2}|\mathcal{F}_{t-1}) = C(F_1(y_{t,1}|\mathcal{F}_{t-1}), F_2(y_{t,2}|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}), \tag{B.4}$$

where C denotes the copula function, $F_1(\cdot|\mathcal{F}_{t-1})$ and $F_2(\cdot|\mathcal{F}_{t-1})$ are the marginal distribution functions of $Y_{t,1}$ and $Y_{t,2}$, respectively, and $\mathcal{F}_{t-1}$ is the filtration.

Having established marginal models in the previous section, we are now left with the construction of the copula C. To this end, we apply the probability integral transform to the standardized residuals to obtain

$$\hat{u}_t^{DE} = F^{DE}(\hat{\eta}_t^{DE}|\mathcal{F}_{t-1}),$$
$$\hat{u}_t^{i} = F^{i}(\hat{\eta}_t^{i}|\mathcal{F}_{t-1}),$$

for $i = 1, \ldots, 31$. F^i and F^{DE} are distribution functions corresponding to each wind site and the German wind index, respectively. Fig. B.6 presents the estimated uniforms for the four selected sites plotted against $\hat{u}_t^{DE}$, showing variations in the strength of the dependence, which we already expected considering the results in Fig. B.1(b). However, Fig. B.6 also indicates that the shapes of the dependence structure could possibly differ, and thus the preferred copula specifications might vary across the wind sites. Specifically, notice that Figs. B.6(b) and B.6(d) exhibit a slight asymmetric behavior with more concentration in the north east corner compared to the other corners. This pattern is not as clear in Figs. B.6(a) and B.6(c).

The copula models we consider are the Gaussian, Gumbel, rotated Gumbel (RGumbel), Clayton, rotated Clayton, Frank, symmetrized Joe-Clayton, and Student t[2]. These are fitted to all 31 pairs and the Akaike Information Criterion (AIC) is used as model selection criterion. To introduce further flexibility, we also consider copula mixtures as in Rodriguez (2007) and Dias and Embrechts (2009). A copula mixture is given by

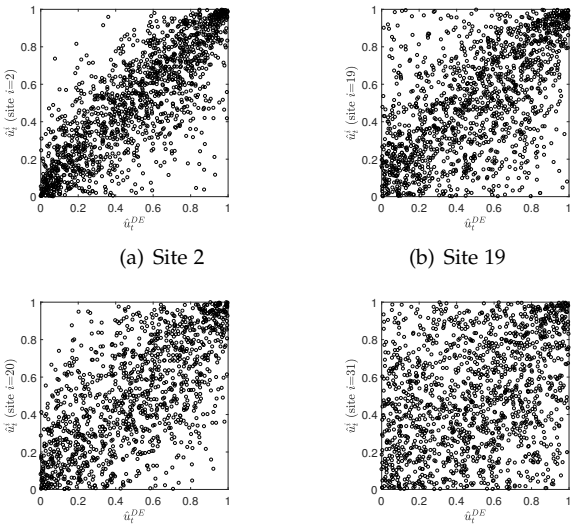
$$c^m(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b, \lambda) = \lambda c^a(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a) + (1-\lambda) c^b(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^b), \tag{B.5}$$

where $c^a(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a)$ is the copula density for copula a having parameter vector $\boldsymbol{\theta}^a$, and $c^b(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^b)$ is the copula density for copula b with parameter vector $\boldsymbol{\theta}^b$. The mixing parameter $0 < \lambda < 1$ controls the proportion of each copula.

[2]Rotation refers to a 180 degree rotation of data.

Site ID	Optimal copula	Kendall's τ	Upper tail dep.	Lower tail dep.	K-S test	CvM test
1	Mix of Gaussian and Gumbel	0.582 (0.011)	0.246 (0.060)	-	0.893	0.886
2	Mix of Gumbel and RGumbel	0.606 (0.008)	0.358 (0.038)	0.328 (0.039)	0.512	0.527
3	Mix of Gaussian and Gumbel	0.534 (0.010)	0.004 (0.044)	-	0.864	0.989
4	Student t	0.642 (0.010)	0.215 (0.123)	0.215 (0.123)	0.968	0.998
5	Mix of Gaussian and Gumbel	0.603 (0.009)	0.074 (0.059)	-	0.809	0.812
6	Mix of Gaussian and Gumbel	0.516 (0.010)	0.189 (0.057)	-	0.902	0.955
7	Mix of Gaussian and Gumbel	0.576 (0.009)	0.056 (0.050)	-	0.857	0.818
8	Mix of Gaussian and Gumbel	0.609 (0.008)	0.082 (0.054)	-	0.881	0.877
9	Mix of Gaussian and Gumbel	0.647 (0.008)	0.333 (0.069)	-	0.783	0.793
10	Mix of Frank and Gumbel	0.453 (0.012)	0.211 (0.048)	-	0.851	0.755
11	Mix of Frank and Gumbel	0.492 (0.011)	0.146 (0.047)	-	0.900	0.765
12	Mix of Gaussian and Gumbel	0.659 (0.007)	0.239 (0.066)	-	0.683	0.789
13	Mix of Gaussian and Gumbel	0.613 (0.009)	0.261 (0.068)	-	0.590	0.642
14	Mix of Frank and Gumbel	0.554 (0.011)	0.373 (0.048)	-	0.683	0.823
15	Mix of Frank and Gumbel	0.570 (0.010)	0.376 (0.046)	-	0.510	0.582
16	Mix of Frank and Gumbel	0.513 (0.011)	0.494 (0.043)	-	0.716	0.797
17	Mix of Frank and Gumbel	0.479 (0.012)	0.395 (0.043)	-	0.358	0.533
18	Mix of Frank and Gumbel	0.570 (0.010)	0.246 (0.050)	-	0.876	0.783
19	Mix of Frank and Gumbel	0.437 (0.013)	0.363 (0.045)	-	0.765	0.703
20	Mix of Frank and Gumbel	0.470 (0.013)	0.190 (0.052)	-	0.958	0.911
21	Mix of Gaussian and Gumbel	0.385 (0.015)	0.223 (0.058)	-	0.982	0.972
22	Mix of Frank and Gumbel	0.506 (0.011)	0.302 (0.048)	-	0.796	0.757
23	Mix of Frank and Gumbel	0.510 (0.013)	0.302 (0.048)	-	0.742	0.864
24	Mix of Frank and Gumbel	0.416 (0.014)	0.196 (0.045)	-	0.504	0.664
25	Mix of Gaussian and Gumbel	0.383 (0.014)	0.271 (0.068)	-	0.800	0.880
26	Mix of Frank and Gumbel	0.437 (0.013)	0.250 (0.048)	-	0.773	0.781
27	Mix of Gaussian and Gumbel	0.349 (0.013)	0.260 (0.059)	-	0.763	0.769
28	Mix of Frank and Gumbel	0.390 (0.014)	0.298 (0.046)	-	0.457	0.558
29	Mix of Frank and Gumbel	0.359 (0.015)	0.186 (0.046)	-	0.696	0.776
30	Mix of Frank and Gumbel	0.361 (0.014)	0.282 (0.046)	-	0.766	0.758
31	Gumbel	0.215 (0.015)	0.277 (0.018)	-	0.756	0.778

Table B.2: Optimal copula, Kendall's τ, upper and lower tail dependence implied by the optimal copula, and p-values resulting from performing goodness-of-fit tests. Standard errors are reported in parenthesis. The symbol "-" indicates that the tail dependence imposed by the chosen copula is zero. Kendall's τ for a convex combination of copula a and copula b, individually imposing a Kendall's τ of τ^a and τ^b, respectively, is $\lambda\tau^a + (1-\lambda)\tau^b$. Likewise, the upper (lower) tail dependence is a convex combination of the individual upper (lower) tail dependence coefficients. See Christensen *et al.* (2019) for more information.



(a) Site 2 (b) Site 19

(c) Site 20 (d) Site 31

Fig. B.6: Empirical copula density plots for four selected sites, with site ID 2, 19, 20, and 31.

In addition to the already mentioned copulas, we employ the following mixtures: Gumbel/RGumbel, Gaussian/Gumbel, Gaussian/RGumbel, Frank/Gumbel, and Frank/RGumbel. For each site, the optimal copula model is reported in Table B.2. Moreover, we include in Table B.2 the implied Kendall's τ, upper tail dependence, and lower tail dependence.[3] The optimal copula differs across the 31 wind sites, with Gaussian/Gumbel and Frank/Gumbel being preferred in most cases. These copula mixtures are both characterized by an asymmetric dependence structure and upper tail dependence. According to the results in Table B.2, only two models corresponding to northern sites imply lower tail dependence, while upper tail dependence is a common feature across all models. The upper tail dependence corresponds to scenarios with high German wind index and high site-specific wind index. For the wind power generator, this feature weakens the power of WPF as hedging instruments, which we shall elaborate in the next section.

To test whether or not the optimal copulas in Table B.2 are well-specified, we perform the Kolmogorov-Smirnov (K-S) and the Cramer-von Mises (CvM) goodness-of-fit tests (see Berg (2009), Genest *et al.* (2009) and Patton (2013) for a detailed description). The resulting p-values are reported in the last two

[3]Recall that tail dependence measures the dependence in extreme events. For positively related data, the upper tail dependence is defined as $\tau^U = \lim_{q \to 1^-} P(u_{1,t} > q | u_{2,t} > q)$, where q denotes the quantile and the u's represent standard uniform variables. The lower tail dependence is defined analogously.

columns of Table B.2 and are based on 999 simulations, indicating that the null hypothesis of a well-specified copula cannot be rejected.

4 Application Results

In this section, the models presented in Sec. 3 are employed to asses the hedging effectiveness of WPF for the different wind sites. As in Sec. 3, we limit ourselves to presenting results for four chosen sites; if, at any point, we do otherwise, this will be clear from the context. In order to perform the hedging exercise, we assume the following: Wind power producers receive a fixed price of 30 EUR per produced MWh regardless of geographical location, and the installed capacity of the wind turbines at each wind site is 100 MW. Thus, from Eq. (B.1) we get that the daily wind power generation at site i and day t is

$$P_{t,i} = 100 \cdot W_{t,i} \cdot h_t,$$

and hence the wind power producer's unhedged profit at site i, over the period $[T_1, T_2]$, is given by

$$R_i^{\mathrm{U}} = 30 \sum_{t=T_1}^{T_2} P_{t,i}. \tag{B.6}$$

When hedging the volume risk associated with wind power generation using a position $\gamma_i \in \mathbb{Z}$ in WPF – recall Eq. (B.2) – we get the following hedged profit

$$R_i^{\mathrm{H}} = R_i^{\mathrm{U}} + \gamma_i R^{\mathrm{WPF}}. \tag{B.7}$$

To determine γ_i, we use 20,000 Monte Carlo simulations from the proposed models, each spanning a year from 1 January 2016. Since the end of the in-sample data is 31 December 2015, it follows that the simulations are performed out-of-sample. We shall restrict our attention to WPF with monthly delivery, and divide therefore each simulated path into 12 parts corresponding to the length of each month. This way, we are able to asses monthly differences in the hedging effectiveness of WPF across a calender year.

Regarding pricing of WPF, i.e., estimating W_{t_0} in Eq. (B.2) for each month, we note that t_0 is fixed at 31 December 2015 regardless of which monthly delivery we consider. Assuming a zero market price of risk, we obtain W_{t_0} as a simple average of Monte Carlo simulations covering the specified delivery period. We stress that this assumption implies that the expected value of the profit distribution will not be affected by changing γ_i. Thus, we concentrate on a minimum variance hedge, and determine γ_i for each month by

$$\min_{\gamma_i \in \mathbb{Z}} \mathbb{V}\mathrm{ar}[R_i^H]. \tag{B.8}$$

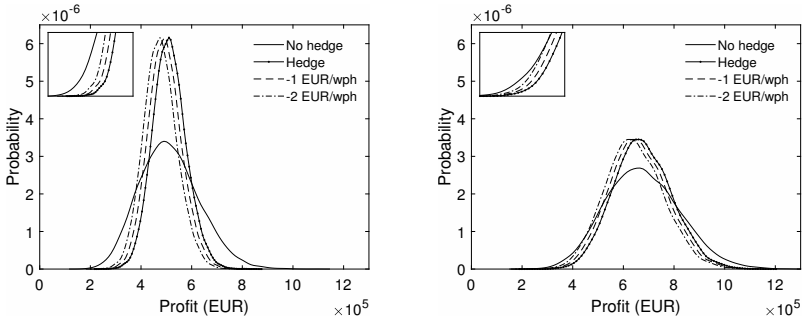
4.1 Hedging Effectiveness of Wind Power Futures

Using the specifications above, we estimate the unhedged profit distributions for all sites and all months of the year. In Fig. B.7, we present the results for January and sites 2, 19, 20, and 31. Note that we are able to compare the profits, since the wind generation capacity is set at 100 MW for all sites. The unhedged profit distributions have different means, with site 31 having the lowest value. Regarding the hedged profit distributions, which are also included in Fig. B.7, we observe that WPF are most beneficial for site 2, followed by sites 19 and 20, which achieve similar benefits. Hedging benefits for site 31 are lowest, since this site is least correlated to the German wind index cf. Table B.2. Although Fig. B.7 only illustrates the case for January, we find similar results for the remaining months. In fact, the average difference between the maximum and minimum variance reductions throughout the year for all sites is only 2.3%, and all range between 1% and 5%.

Motivated by the earlier studies of Gersema and Wozabal (2017) and Benth and Pircalabu (2017), who argue for the existence of a negative risk premium in the German wind power futures market, we consider next the effect of including such a quantity in our analysis. Cf. Eq. (B.2), a risk premium of e.g. -1 EUR/wph translates to a reduction in W_{t_0} of 1 %, and implies that WPF are sold at a discount. To indicate the effect of a negative risk premium on the hedged profit distribution, we plot in Fig. B.7 the results of assuming a risk premium of -1 EUR/wph and -2 EUR/wph. We note that the hedge quantities remain unchanged, since hedging decisions are based on a minimum variance hedge, thus disregarding the cost of the hedge portfolio. In the hedging exercise considered here, this implies that a negative risk premium shifts the hedged profit distribution to the left compared to the case of a zero risk premium, as it appears from Fig. B.7.[4]

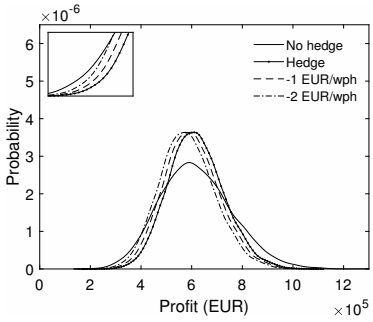
To study further the implications of a negative risk premium, we consider the reduction in the mean of the hedged profit distribution when taking into account a risk premium of -2 EUR/wph, as opposed to none at all. The results are plotted in Fig. B.8(a) on a monthly basis, and can be explained by the seasonality in the marginal models in combination with a static copula. On one hand, the seasonal behavior embedded in the site-specific wind index causes the profit corresponding to the naked position R_i^U to be lowest during summer and highest during winter. On the other hand, the optimal hedge quantity γ^* in WPF is highly influenced by the dependence – which is constant throughout the year – thus translating to a fairly constant γ^* as illustrated in Fig. B.8(b).

[4]In practice, the hedger is not indifferent to the cost of the hedging portfolio. The objective in Eq. (B.8) can easily be extended to e.g. a situation where the variance of the profit distribution is minimized while also minimizing the cost of the hedging portfolio, based on some preference regarding the trade-off between variance and cost reduction. This is however outside the scope

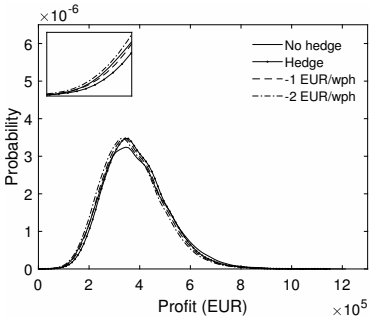


(a) Site 2

(b) Site 19

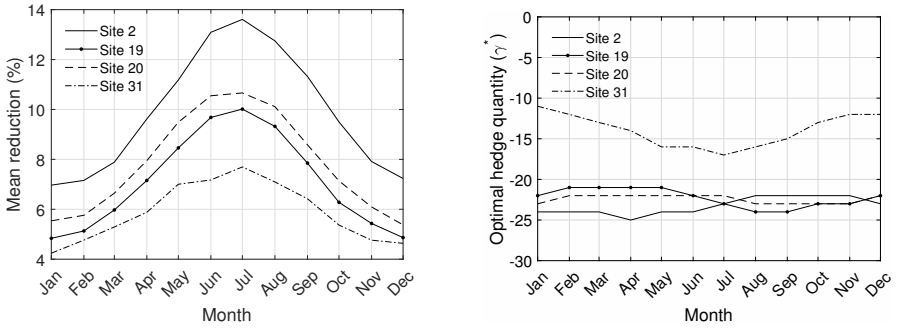


(c) Site 20



(d) Site 31

Fig. B.7: Unhedged and hedged profit distributions for wind sites 2, 19, 20 and 31 cf. Fig. B.1(a) for January. The results are obtained by employing the marginal models determined in Sec. 3.1 and the optimal copulas cf. Table B.2. The hedged distributions are computed with a risk premium corresponding to 0 EUR/wph, -1 EUR/wph and -2 EUR/wph. The optimal hedge quantities are obtained by solving Eq. (B.8), and do not vary with the risk premium.



(a) Percentage reduction in mean (b) Positions in WPF

Fig. B.8: Percentage reduction in mean of the hedged profit distribution when including a risk premium of -2 EUR/wph as opposed to no risk premium, and the optimal number of WPF contracts for each month throughout the year. The results are displayed for wind sites 2, 19, 20, and 31 cf. Fig. B.1(a).

Clearly, since wind power generators seek to cover their exposure to low wind scenarios, they should be more inclined to use WPF as hedging instruments during the summer months. According to the results in Fig. B.8(a) however, the reduction in the mean of the hedged profit distribution is approximately double as high during summer compared to winter, which is highly unfavorable. We stress that the results in Fig. B.8(a) only hold under the assumption of a constant risk premium across the months of a calendar year, which cannot be the case in practice. Nevertheless, since the empirical investigation carried out in Benth and Pircalabu (2017) does not suggest any clear seasonal pattern in the market price of risk associated with German WPF, we argue that our results are relevant from a practical perspective. Furthermore, the analysis above would only be rendered superfluous in a situation where the risk premium was shaped as to counteract the yearly seasonality in the wind index.

On a different note, we mention in passing that the yearly seasonality associated with wind power production is also important in the context of liquidity in the German WPF market. According to Christensen *et al.* (2019), hedging benefits for conventional power producers are highest during winter. Hence, a large part of the buyer side is incentivized to take a position in WPF during winter, whereas the seller side is incentivized to do so during summer.

Next, we consider the variance reduction achieved by hedging with WPF. In Fig. B.9(a), we present the estimated average variance reductions throughout the year. As expected, the variance reductions and geographical location of wind turbines are related, with the reductions decreasing as we move from

of the present paper, and shall not be pursued further.

north to south Germany. The results vary substantially across the wind sites, with the highest numbers being above 70%, and the lowest number being 13%. Also relevant to consider in this context is the effect of variance reduction on the tails of the hedged profit distribution. Under the assumption of a zero risk premium, we wish to quantify whether the right tail is reduced by more than the left tail, i.e., are we reducing the probability of a very high income by more than that of a very low income? To answer this question, we let $q^{\mathrm{H}}_{j,i}$ denote the jth quantile of the demeaned hedged profit distribution for site i,

$$q^{\mathrm{H}}_{j,i} = F^{-1}(R^{\mathrm{H}}_i - \mathbb{E}[R^{\mathrm{H}}_i], j).$$

Similarly, $q^{\mathrm{U}}_{j,i}$ denotes the jth quantile of the demeaned unhedged profit distribution. The percentage change of the jth quantile from performing the WPF hedge can then be measured by

$$T^{\mathrm{l}}_i(j) = \frac{q^{\mathrm{U}}_{j,i} - q^{\mathrm{H}}_{j,i}}{q^{\mathrm{U}}_{j,i}}.$$

To measure the difference between the change in the tails, we define the variable,

$$T_i(j) = |T^{\mathrm{l}}_i(j)| - |T^{\mathrm{r}}_i(j)|,$$

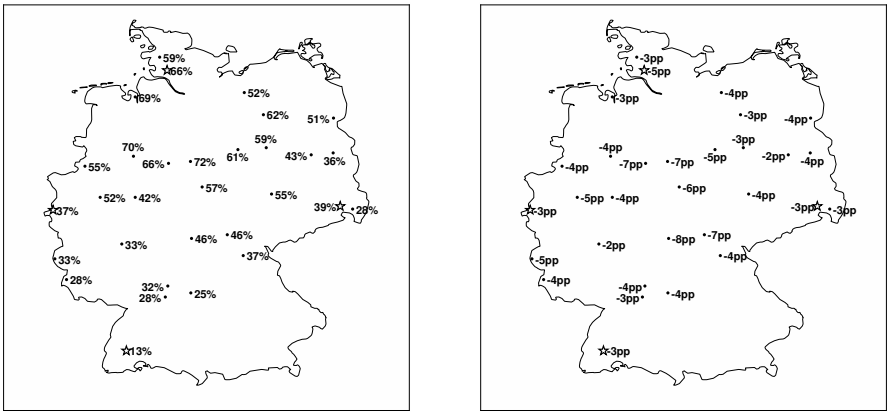
where

$$T^{\mathrm{r}}_i(j) = \frac{q^{\mathrm{H}}_{1-j,i} - q^{\mathrm{U}}_{1-j,i}}{q^{\mathrm{U}}_{1-j,i}}.$$

The quantity $T^{\mathrm{l}}_i(j)$ defined above will be positive if we shrink the left tail (jth quantile) when performing the WPF hedge, which is the case for all sites. Further, $T^{\mathrm{r}}_i(j)$ will be negative if we shrink the right tail ($1-j$th quantile) when performing the WPF hedge, which is also the case for all sites. Hence, a negative value of $T_i(j)$ means that the right tail has shrunk more than the left tail. We note that $T_i(j)$ is measured in percentage points (pp). Estimates for $T_i(0.05)$ are reported in Fig. B.9(b), revealing negative values for all sites, with -8pp being the lowest. Interestingly, the magnitude of the average variance reduction does not seem to influence the magnitude of $T_i(0.05)$, cf. Figs. B.9(a) - B.9(b). For example, the lowest value of $T_i(0.05)$ corresponds to a variance reduction of 46%, which is significantly lower than the highest variance reduction of 72%.

Since a zero risk premium is imposed due to the results in Fig. B.9(b), the negative values of $T_i(0.05)$ obtained for all sites can be explained by 1)

the marginal specifications and 2) the presence of asymmetry and tail dependence in the copula. On one hand, the marginal distributions of wind indexes are skewed, with a heavy tail to the right. On the other hand, cf. Table B.2, evidence for asymmetry and upper tail dependence is found for most wind sites, entailing that site-specific wind indexes are related to the German index in extreme high wind scenarios. Since wind power producers seek to cover their exposure to low wind scenarios, lower tail dependence would be much preferred to the upper tail dependence that we find in the data. It is however difficult to separate the effects related to the margins from those related to the dependence structure without further analysis. In the next section, we return to this issue. Lastly, we stress that the results in Fig. B.9(b) are obtained assuming a zero risk premium. In the presence of a negative risk premium, the negative effects illustrated in Fig. B.9(b) will clearly be magnified.



(a) Average variance reductions (b) $T_i(0.05)$

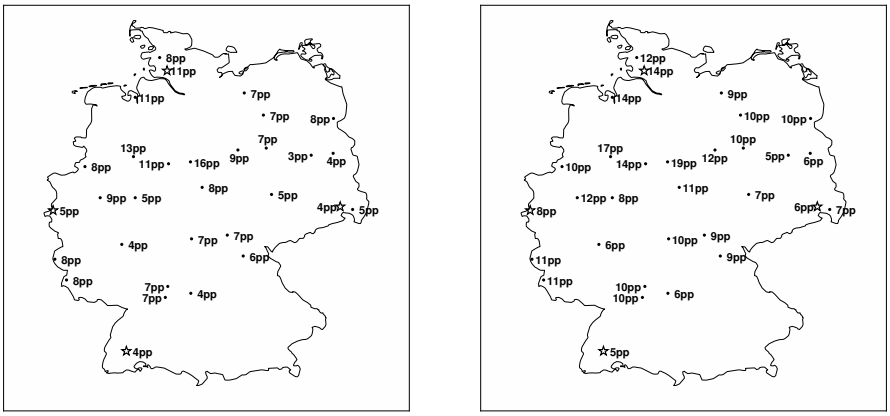
Fig. B.9: Average variance reduction for each wind site using the optimal copula cf. Table B.2, and the difference between the change in the tails of the profit distribution measured by $T_i(0.05)$. The stars indicate the location of the sites with ID 2, 19, 20, and 31.

4.2 Comparison Study

To highlight the effects of employing a copula that allows for asymmetry and upper tail dependence, we consider in this section a comparison study. Specifically, we rerun the computations performed in Section 4.1 with the Frank copula, while keeping the marginal models fixed. We employ the Frank copula as a naive alternative since Frank is represented in the optimal copula mixtures in 16 out of 31 cases cf. Table B.2. In addition, the Frank copula is symmetric and imposes zero tail dependence, whereas the optimal

copulas are all asymmetric (except for site 4) and imply tail dependence (upper tail dependence for the most part). Further, by performing the K-S and CvM goodness-of-fit tests, we cannot reject the Frank copula for any of the 31 sites at a 5% significance level. In fact, if we excluded the copula mixtures from our analysis, Frank would become the preferred copula specification for many wind sites based on AIC. In the absence of copula mixtures, a hedger or risk manager could therefore easily have employed a Frank copula to model dependence.

In Fig. B.10(a), we report the average difference in the percentage variance reduction across the months of the year, when going from the optimal copula cf. Table B.2 to the Frank copula. We identify increases in the average variance reduction for all sites when using the optimal copula. We complement these findings with Fig. B.10(b), where we report the maximum monthly difference between percentage variance reduction implied by the optimal copula and Frank. The maximum difference reaches as high as 19pp, revealing that using a less optimal model can have a substantial impact. In a nutshell, we would simply undermine the variance reduction strength of WPF with the Frank copula.



(a) Average difference (b) Maximum difference

Fig. B.10: Average and maximum difference in variance reductions implied by the optimal copula cf. Table B.2 and the Frank copula (optimal minus naive). The maximum differences refer to the largest monthly difference. The stars indicate the location of the sites with ID 2, 19, 20, and 31.

Finally, recall Fig. B.9(b), where it was difficult to separate the effects from the marginals and the dependence structure. Performing similar calculations as in the case of Fig. B.9(b) with the Frank copula yields the estimates of $T_i(0.05)$ given in Fig. B.11. Compared to Fig. B.9(b), we find that introducing

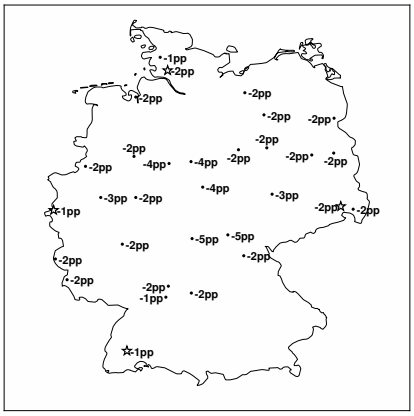


Fig. B.11: Estimates of $T_i(0.05)$ when performing the hedge with WPF using the naive approach, i.e, the Frank copula. The stars indicate the location of the sites with ID 2, 19, 20, and 31.

a different dependence structure accounts for approximately 2-3pp of the difference between left and right tail percentage change. Hence, using the Frank copula would lead to an underestimation of the difference between the change in tails implied by hedging with WPF.

5 Conclusion

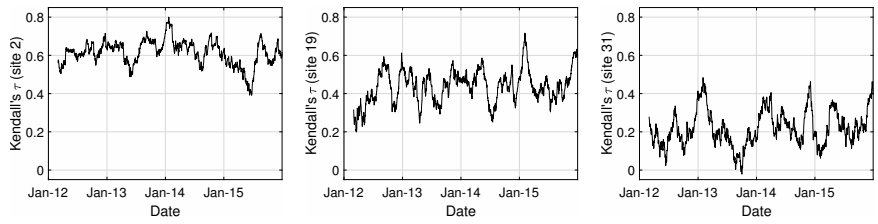
In this paper, we analyze the hedging effectiveness of wind power futures for wind power generators with 31 different locations in Germany. We propose ARMA–GARCH copula models for each German wind index and site-specific wind index pair. The ARMA–GARCH type models capture the marginal behavior of the variables rather well, and a detailed study of the different dependence structures reveals a general tendency of asymmetric dependence and upper tail dependence, thus justifying the use of copulas.

Based on Monte Carlo simulations from the proposed models, the benefits of wind power futures are quantified through a minimum variance hedge, which we apply for each wind site. The variance reductions vary from 13% to 72%, with the lowest values corresponding to wind farms located in the south of Germany. Motivated by earlier findings in the literature that argue for the existence of a negative risk premium in the German wind power futures market, we extend our analysis to include this feature. Compared to the case of a zero risk premium, the hedged profit distribution is shifted to the left, thus weakening the benefits of wind power futures in the context of our paper. Considering how the wind power futures hedge affects the tails of the profit distributions, our results indicate a higher decrease in the upside potential compared to the decrease in the downside risk. While this is

expected due to the presence of skewness in the unhedged profit distribution, the asymmetric dependence plays an important role as well.

Lastly, we highlight the importance of capturing asymmetry and upper tail dependence by performing a comparison study, where the hedging exercise is carried out assuming a naive model for the dependence structure. With the naive specification, the variance reduction resulting from the wind power futures hedge would be undermined. Further, approximately half of the difference in reduction between the tails in the hedged profit distribution is caused by the asymmetric dependence.

As a concluding remark, we shortly address the topic of time-varying dependence. Fig. B.12 shows Kendall's τ between the German wind index and three site-specific wind indexes (site 2, 19, and 31) based on a 60-days moving window. Not surprisingly, Fig. B.12 reveals that the dependencies vary through time; however, they do stay in a rather narrow range (this applies for all sites). Like in Grothe and Schneiders (2011), we have neglected this time variation by restricting our attention to constant copula models. Nevertheless, extending a copula model to capture time-varying behavior can be achieved through e.g. the GAS model proposed by Creal *et al.* (2013). While such extension is interesting, analysing the impact of time-varying dependencies on hedging benefits is left as future research.



(a) Site ID 2 (b) Site ID 19 (c) Site ID 31

Fig. B.12: Time series plots of Kendall's τ between three site-specific wind indexes and the German wind index based on a 60-days moving window.

Acknowledgements

The authors would like to thank Esben Høg and the Quantitative Analytics team at Centrica Energy Trading for valuable discussions. One anonymous referee is also thanked for her constructive criticism and suggestions, which improved the presentation of this paper.

Declarations of Interest

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B. Anca Pircalabu is supported by the Innovation Fund Denmark under Grant 4135-00082B.

References

Aepli, M. D., Füss, R., Henriksen, T. E. S. and Paraschiv, F. (2017). Modeling the multivariate dynamic dependence structure of commodity futures portfolios. *Journal of Commodity Markets*, **6**, 66–87. https://doi.org/10.1016/j.jcomm.2017.05.002

Avdulaj, K. and Barunikl, J. (2015). Are benefits from oil - stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, **51**, 31–44. https://doi.org/10.1016/j.eneco.2015.05.018

Benth, F. E. and Šaltytė Benth, J. (2011). Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, Vol. 2011, Article ID 576791, 21 pages. http://dx.doi.org/10.1155/2011/576791

Benth, F.E. and Kettler, P. C. (2011). Dynamic copula models for the spark spread. *Quantitative Finance*, **3**, 407–421. https://doi.org/10.1080/14697688.2010.481629

Benth, F. E. and Pircalabu, A. (2017). A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures. *Working paper*. Available at SSRN: https://ssrn.com/abstract=2979341.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *European Journal of Finance*, **15**(7–8), 675–701. https://doi.org/10.1080/13518470802697428

Börger, R., Cartea, A., Kiesel, R. and Schindlmayr, G. (2009). Cross-commodity analysis and applications to risk management. *Journal of Futures Markets*, **29**(3), 197–217. https://doi.org/10.1002/fut.20359

Christensen, T. S., Pircalabu, A., and Høg, E. (2019). A seasonal copula mixture for hedging the clean spark spread with wind power futures. *Energy Economics*, **78**, 64–80.

Creal, D., Koopman, S.J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, **28**, 777-795. https://doi.org/10.1002/jae.1279

Dias, A. and Embrechts, P. (2009). Testing for structural changes in exchange rates' dependence beyond linear correlation. *The European Journal of Finance*, **15**(7), 619–637. https://doi.org/10.1080/13518470701705579

Elberg, C. and Hagspiel, S. (2015). Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, **241**, 260–272. https://doi.org/10.1016/j.ejor.2014.08.026

Genest, C., Remillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, **44**(2), 199–213. https://doi.org/10.1016/j.insmatheco.2007.10.005

Gersema, G. and Wozabal, D. (2017). An equilibrium pricing model for wind power futures. *Energy Economics*, **65**, 64–74. https://doi.org/10.1016/j.eneco.2017.04.032

Grothe, O. and Schneiders, J. (2011). Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, **39**, 4742–4754. https://doi.org/10.1016/j.enpol.2011.06.052

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, **94**(2), 401–419. https://doi.org/10.1016/j.jmva.2004.06.003

Liu, B.-Y., Ji, Q. and Fan, Y. (2017). Dynamic return-volatility dependence and risk measure of CoVaR in the oil market: A time-varying mixed copula model. *Energy Economics*, **68**, 53–65. https://doi.org/10.1016/j.eneco.2017.09.011

McNeil, A. J., Frey, R. and Embrechts, P. (2005). Quantitative risk management. *Princeton Series in Finance*.

Nelsen, R.B. (2006). *An Introduction to Copulas*, Springer.

Patton, A.J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, **47**(2), 527–556. https://doi.org/10.1111/j.1468-2354.2006.00387.x

Patton, A.J. (2013). Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899–960. Elsevier B.V. https://doi.org/10.1016/B978-0-444-62731-5.00016-6

Pircalabu, A. and Jung, J. (2017). A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quantitative Finance*, **17**(10), 1583–1600. https://doi.org/10.1080/14697688.2017.1307511

Pircalabu, A., Hvolby, T., Jung, J. and Høg, E. (2017). Joint price and volumetric risk in wind power trading: A copula approach. *Energy Economics*, **62**, 139–154. https://doi.org/10.1016/j.eneco.2016.11.023

Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance*, **14**(3), 401–423. https://doi.org/10.1016/j.jempfin.2006.07.002

Šaltytė Benth, J. and Benth, F. E. (2010). Analysis and modelling of wind speed in New York. *Journal of Applied Statistics*, **37**, 893–909. https://doi.org/10.1080/02664760902914490

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, **8**, 229–231.

Tol, R.S.J. Autoregressive Conditional Heteroscedasticity in daily wind speed measurements. *Theoretical and Applied Climatology*, **56**, 113–122. https://doi.org/10.1007/BF00863788

Paper C

Multivariate Continuous-Time Modeling of Wind Indexes and Hedging of Wind Risk

List of authors: Fred Espen Benth[1], Troels Sønderby Christensen[2,3], Victor Rohde[4]

[1]Department of Mathematics, University of Oslo
[2]Department of Mathematical Sciences, Aalborg University
[3]Quantitative Analytics, Centrica Energy Trading
[4]Department of Mathematics, Aarhus University

The paper has been submitted to Quantitative Finance.

The layout has been revised.

Abstract

With the introduction of the exchange-traded German wind power futures, opportunities for German wind power producers to hedge their volumetric risk are present. We propose two continuous-time multivariate models for the wind power utilization at different wind sites, and discuss the properties and estimation procedures for the models. Employing the models to wind index data for wind sites in Germany and the underlying wind index of exchange-traded wind power futures contracts, the estimation results of both models suggest that they capture key statistical features of the data. We argue how these models can be used to find optimal hedging strategies using exchange-traded wind power futures for the owner of a portfolio of so-called tailor-made wind power futures. Both in-sample and out-of-sample hedging scenarios are considered, and, in both cases, significant variance reductions are achieved. Additionally, the risk premium of the German wind power futures is analysed, leading to an indication of the risk premium of tailor-made wind power futures.

1 Introduction

In the power market, producers in general face market risk in the sense of uncertainty of the prices at which they can sell their generated power. The intermittent nature of renewable energy sources such as wind and photo voltaic power production adds yet another layer of risk, known as volumetric risk in the sense that the produced amount of electricity is uncertain due to the dependence on weather. Globally, so-called power purchase agreements and subsidies from governments have minimized the market risk for renewable power producers. In contrast, the volumetric risk has only recently been addressed in Germany—and only for wind power producers (WPPs)—by the introduction of the exchange-traded wind power futures (WPF) contracts. The underlying of a WPF contract is an index between zero and one representing the overall utilization of the installed German wind power production. By taking an appropriate position in WPF contracts, the lost income of the WPPs implied by low wind scenarios is (partially) offset by the position in WPF contracts, hence minimizing the volumetric risk. Due to the prioritization of the cheapest power producers in Germany, the opposite part of the WPF market is typically conventional power producers (CPPs) such as gas-fired power plants. By taking an appropriate position in exchange-traded WPF contracts, CPPs can hedge their exposure to the cheap electricity generated by WPPs.

The WPF market is considered in detail in Gersema & Wozabal (2017), where the authors propose an equilibrium pricing model. They find that the willingness to engage in the WPF market is greater for the WPPs compared to CPPs. In other words, the hedging benefits of the WPF contracts are

greater for WPPs than CPPs. This is supported by the results in Christensen & Pircalabu (2018) who employ an ARMA-GARCH copula framework to the joint modelling of one site-specific wind index and the underlying WPF index. In Bertsekas & Pircalabu (2018) modeling of the underlying WPF index is considered and closed-form formulas for the WPF price and the price of European options written on the WPF index are derived.

Continuous-time modeling using univariate Ornstein-Uhlenbeck (OU) processes driven by non-decreasing Lévy processes, like the compound Poisson process with exponential jumps, have been studied extensively, and used to model, for example, stochastic volatility of financial assets, wind, electicity prices, and temperature (see Barndorff-Nielsen & Shephard (2001); Benth & Benth (2007); Benth *et al.* (2007); Bertsekas & Pircalabu (2018)). A detailed treatment of Lévy processes can be found in Sato (1999). The multivariate modeling of more than two stochastic processes using multidimensional non-Gaussian Lévy processes is, however, more limited. Here we mention the work of Leoni & Schoutens (2008) and Semeraro (2008) that introduce the multivariate construction by subordination of Brownian motions, and the work of Ballotta & Bonfiglioli (2016) using linear transformations of Lévy processes.

Our contribution to the literature is twofold. Firstly, we propose a joint model for the simultaneous behaviour of wind indexes that allows for a parsimonious representation of the correlation structure. This model can be seen as the multivariate version of the model presented in Bertsekas & Pircalabu (2018). As a consequence of the scarce literature on such multivariate models, we propose an alternative model for comparison reasons. Secondly, we suggest the idea of so-called tailor-made WPF contracts to eliminate the volumetric risk of WPPs completely. Employing our proposed joint model of wind indexes, we investigate the hedging benefits of exchange-traded WPF contracts for a owner of a portfolio of tailor-made WPF contracts, and comment on the risk premium of tailor-made WPF contracts. We show that this construction is beneficial for both parties of the tailor-made WPF contracts.

The rest of the paper is organized as follows. Sec. 2 presents the data of wind indexes we later analyze in greater detail and also serves as motivation for the proposed model. In Sec. 3 we present models for the joint behaviour of wind indexes and corresponding estimation procedures. In Sec. 4 we present the estimation results of the two models. Sec. 5 discusses the hedging of wind power production using WPF contracts implied by the proposed models. Lastly, Sec. 6 concludes.

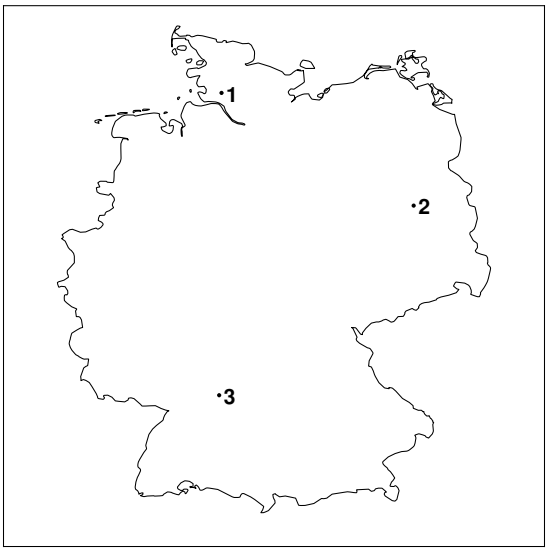


Fig. C.1: Locations of wind sites with site ID in Germany.

2 Data Presentation

The empirical observation period spans from 1 July 2016 to 30 June 2019, which corresponds to 1095 daily observations for each considered wind index. The data set consists of

1. A daily wind index at three wind sites provided by Centrica Energy Trading. The wind index at wind site i is calculated by

$$\frac{Q_i(t)}{h(t)C_i},$$

where $h(t)$ is the number of hours in day t, $Q_i(t)$ is the power production at day t at site i, and C_i is the installed capacity at site i. Figure C.1 shows the approximate geographical locations for the three wind sites.[1]

2. A daily German wind index provided by Nasdaq, representing the German utilization of wind power plants. The acronym used for it is NAREX WIDE (NAsdaq REnewable indeX WInd DE (Germany)) and is used as the underlying for WPF contracts traded on Nasdaq. We will simply denote it as the German wind index in the remaining part of the paper.

[1]The locations are approximate due to confidentially issues.

The wind indexes are bounded between zero and one. Fig. C.2 shows all four wind indexes, and the corresponding autocorrelation function for each wind index. In all four cases, the wind index displays a pronounced yearly cycle consistent with the observations made in Bertsekas & Pircalabu (2018) and Christensen & Pircalabu (2018). Since the German wind index is by construction made up of all wind power production in Germany, the behaviour of the German wind index is less extreme compared to the individual wind indexes. To concretize, a value of zero for the German wind index is not observed in our observation period, whereas it is for all three wind sites. Also the maximum attained value for each site wind index is higher than the maximum value of the German index; however, it does not reach one in any of the cases.

3 Model Description

3.1 General Model Considerations

Let n denote the number of wind sites and $P_i(t)$ the ith wind index. We assume that the ith wind index can be described by

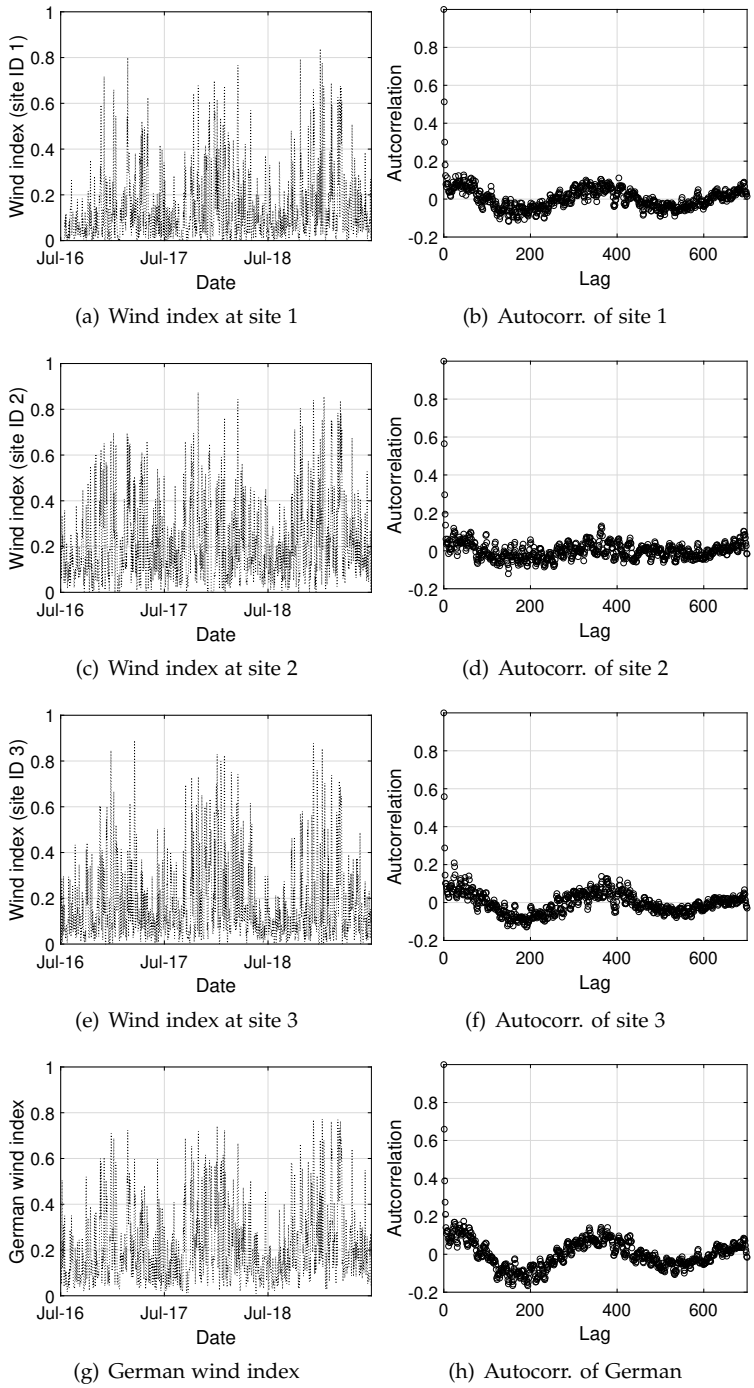
$$P_i(t) = 1 - e^{-S_i(t)X_i(t)}, \quad i = 1, \ldots, n, \tag{C.1}$$

where $S_i(t) : \mathbb{R} \to \mathbb{R}^+$ is a deterministic function intended to filter out potential seasonal effects, and $X_i(t)$ is a mean-reverting stochastic process satisfying $X_i(t) \geq 0$ for all t. The intention of $X(t) = (X_1(t), ..., X_n(t))^\top$ is to capture the short-term uncertainty and the dependence between the n wind indexes. By this specification we are ensured that $P_i(t) \in [0, 1)$.

The proposed model in Eq. (C.1) distinguishes itself from the specification in Bertsekas & Pircalabu (2018), where the natural extension of their univariate setup to the present multivariate setup would be

$$P_i(t) = S_i(t)e^{-X_i(t)}. \tag{C.2}$$

with appropriate choices of $X_i(t)$ and $S_i(t)$. We do, however, prefer Eq. (C.1) over Eq. (C.2). Due to our specification with regard to the deterministic part $S_i(t)$ of the model, we do not face any potential model inconsistencies as is the case of Eq. (C.2). We refer the interested reader to Bertsekas & Pircalabu (2018) more information and discussion. Additionally, as discussed in Sec. 2, the wind index at a given site can attain a value of zero, whereas, on the other hand, we have not observed full utilization of the capacity at a single wind site. Since Bertsekas & Pircalabu (2018) consider the German wind index separately, which is never zero or one due to the construction of it, Eq. (C.2) is applicable without any modifications. Lastly, Eq. (C.1) implies



(a) Wind index at site 1 (b) Autocorr. of site 1

(c) Wind index at site 2 (d) Autocorr. of site 2

(e) Wind index at site 3 (f) Autocorr. of site 3

(g) German wind index (h) Autocorr. of German

Fig. C.2: All four wind indexes with corresponding empirical auto-correlation function.

that increased values of $S_i(t)$ and $X_i(t)$ translate to an increased value of $P_i(t)$, which is more intuitively appealing.

Moving on to the seasonal components of the model, we include the following yearly seasonality motivated by the observations made in Fig. C.2,

$$S_i(t) = a_i + b_i \sin(2\pi t/365) + c_i \cos(2\pi t/365), \quad i = 1, \ldots, n,$$

where $a_i, b_i, c_i \in \mathbb{R}$. With N being the number of observations, the coefficients are determined by

$$\min_{a_i, b_i, c_i} \sum_{t=1}^{N} \left[-\log\left(1 - P_i(t)\right) - S_i(t) \right]^2, \quad i = 1, \ldots, n.$$

Having obtained the estimated seasonal functions, the observed values of $X_i(t)$ implied by the estimated seasonal function $\hat{S}_i(t)$ can then be calculated as

$$X_i(t) = \frac{-\log(1 - P_i(t))}{\hat{S}_i(t)}. \tag{C.3}$$

We will in the following discuss two approaches for modeling $X_i(t)$.

3.2 A Gamma Model

In this section a multivariate model for $n - 1$ wind indexes and the German wind index is discussed, which we refer to as the gamma model in the sequel. In Sec. 5.1 we will consider the case $n = 4$. We start by introducing the noise process. In particular, we say a Lévy process L is a compound Poisson process with exponential jumps and parameters $\alpha > 0$ and $\beta > 0$ if

$$L(t) = \sum_{i=1}^{N(t)} J_i$$

where $(N(t))_{t \in \mathbb{R}}$ is a Poisson process with intensity α and J_i, $i \in \mathbb{N}$, are independent exponentially distributed random variables with parameter β. We say a random variable has an exponential distribution with parameter β if it has density $x \mapsto \mathbb{1}_{[0,\infty)}(x)\beta e^{-\beta x}$.

The gamma model assumes X is a multidimensional Lévy-driven Ornstein-Uhlenbeck (OU) process,

$$dX(t) = -\Lambda X(t)dt + \Sigma_L dL(t). \tag{C.4}$$

Here, L is an n-dimensional Lévy process where the i'th entry is an independent compound Poisson process with exponential jumps, variance equal to one, and parameters α_i and β_i for $i = 1, \ldots, n$. Furthermore, Λ is a diagonal

matrix, $\text{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i > 0$ for $i = 1, \ldots, n$. We assume Σ_L is given by

$$\Sigma_L = \begin{pmatrix} \sigma_{1,1} & 0 & \ldots & 0 & \sigma_{1,n} \\ 0 & \sigma_{2,2} & \ldots & 0 & \sigma_{2,n} \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & \sigma_{n-1,n-1} & \sigma_{n-1,n} \\ 0 & 0 & 0 & 0 & \sigma_{n,n} \end{pmatrix} \tag{C.5}$$

and that all entries of Σ_L are non-negative. Due to the form of Σ_L, each individual wind index has an idiosyncratic risk associated to it through one of the first $n-1$ compound Poisson process $L_1, \ldots, L_{n-1}$. Furthermore, all sites and the German index share a systematic risk through the n'th compound Poisson process L_n. A similar construction is also considered in Ballotta & Bonfiglioli (2016) where a multivariate model is proposed for modeling financial products written on more than one underlying asset.

Distribution of $P_n(t)$ in the Gamma Model

Since X_n, the process associated with the German wind index, is an OU process driven by one compound Poisson process with exponential jumps, it has a gamma distribution as its stationary distribution. Thus, we conclude the following:

Proposition 3.1. *The stationary distribution of $P_n(t)$ in the gamma model has density*

$$f_{P_n(t)}(x) = \frac{(-\log(1-x))^{\alpha_n - 1}(1-x)^{\beta_n / S_n(t) - 1}}{S_n(t)^{\alpha_n}} \quad x \in (0,1) \tag{C.6}$$

Proof. This is a direct consequence of $X_n(t)$ being gamma distributed with shape α_n and rate β_n (see for example Barndorff-Nielsen & Shephard (2001)). □

In Figure C.3 the density of P_n is depicted for different α_n and β_n (with $S_n(t) = 1$). We see that the densities implied by the gamma model are rather flexible and able to cover both low and high utilization scenarios.

The processes $X_1, \ldots, X_{n-1}$ are sums of two independent gamma distributions, and thus, there does not, in general, exist simple expressions for the densities of the individual site index similar to the one for the German index stated in Prop. 3.1.

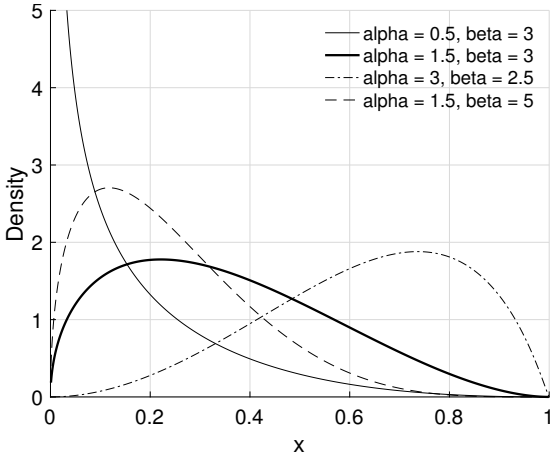


Fig. C.3: Different variations of the density in (C.6)

Covariance Between Wind Indexes in the Gamma Model

We now give (semi-)analytical expressions of the covariances implied by the gamma model, which will be useful for fast calculation of the minimum variance hedges discussed in Section 5. The integral in Eq. (C.7) is the only non-analytical part of the expression, but we argue in Remark A.3 that this integral is small and can be coarsely approximated without significant effect. We thereby maintain fast computation time.

Before we state the result, let us introduce some notation to help making a concise statement. To this end define the $n \times (n+1)$ matrix $\tilde{\Sigma}_L$ by

$$\tilde{\Sigma}_L = \begin{pmatrix} \sigma_{1,1} & 0 & \dots & 0 & 0 & \sigma_{1,n} \\ 0 & \sigma_{2,2} & \dots & 0 & 0 & \sigma_{2,n} \\ \vdots & \vdots & \ddots & 0 & 0 & \vdots \\ 0 & 0 & 0 & \sigma_{n-1,n-1} & 0 & \sigma_{n-1,n} \\ 0 & 0 & 0 & 0 & 0 & \sigma_{n,n} \end{pmatrix}$$

where $\sigma_{i,j}$ is the (i,j)'th entry of Σ_L. Let $\tilde{\sigma}_{i,j}$ denote the (i,j)'th entry of $\tilde{\Sigma}_L$. Furthermore, define $\tilde{\alpha}, \tilde{\beta} \in \mathbb{R}^{n+1}$ by

$$\tilde{\alpha} = (\alpha_1, \dots, \alpha_{n-1}, 0, \alpha_n)^\top \quad \text{and} \quad \tilde{\beta} = (\beta_1, \dots, \beta_{n-1}, 1, \beta_n)^\top,$$

and denote the i'th entry of $\tilde{\alpha}$ and $\tilde{\beta}$ by $\tilde{\alpha}_i$ and $\tilde{\beta}_i$. We now give the expressions of the covariances of the gamma model. The proof is relegated to the Appendix.

Proposition 3.2. *Consider $s \leq t$ and define*

$$f_{i,j}(u) = S_i(t)\tilde{\sigma}_{i,n+1}e^{-\lambda_i(t-s+u)} + S_j(s)\tilde{\sigma}_{j,n+1}e^{-\lambda_j u}, \quad i,j = 1,\dots,n.$$

Then

$$\begin{aligned}
&\operatorname{cov}(P_i(t), P_j(s)) \\
&\quad = \left(\frac{\tilde{\beta}_i}{\tilde{\beta}_i + \tilde{\sigma}_{i,i} S_i(t)}\right)^{\tilde{\alpha}_i/\lambda_i} \left(\frac{\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_i(t) e^{-\lambda_i(t-s)}}{\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_i(t)}\right)^{\tilde{\alpha}_{n+1}/\lambda_i} \\
&\quad \times \left(\frac{\tilde{\beta}_j}{\tilde{\beta}_j + \tilde{\sigma}_{j,j} S_j(s)}\right)^{\tilde{\alpha}_j/\lambda_j} \left[\left(1 + \frac{f_{i,j}(0)}{\tilde{\beta}_{n+1}}\right)^{\tilde{\alpha}_{n+1} f_{i,j}(0)/f'_{i,j}(0)}\right. \\
&\quad \times \exp\left\{\tilde{\alpha}_{n+1} \int_0^\infty \frac{d}{du}\left(\frac{f_{i,j}(u)}{\frac{d}{du} f_{i,j}(u)}\right) \log\left(1 + \frac{f_{i,j}(u)}{\tilde{\beta}_{n+1}}\right) du\right\} \\
&\quad \left. - \left(\frac{\tilde{\beta}_{n+1}}{\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_4(t) e^{-\lambda_i(t-s)}}\right)^{\tilde{\alpha}_{n+1}/\lambda_i} \left(\frac{\tilde{\beta}_{n+1}}{\tilde{\beta}_{n+1} + \tilde{\sigma}_{j,n+1} S_j(s)}\right)^{\tilde{\alpha}_{n+1}/\lambda_j}\right]
\end{aligned} \tag{C.7}$$

for $i, j = 1, \ldots, n$, $i \neq j$, *and*

$$\begin{aligned}
&\operatorname{cov}(P_i(t), P_i(s)) \\
&\quad = \left(\frac{\tilde{\beta} i + \tilde{\sigma}_{i,i} S_i(t) e^{-\lambda_i(t-s)}}{\tilde{\beta}_i + \tilde{\sigma}_{i,i} S_i(t)}\right)^{\tilde{\alpha}_i/\lambda_i} \left(\frac{\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_i(t) e^{-\lambda_i(t-s)}}{\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_i(t)}\right)^{\tilde{\alpha}_{n+1}/\lambda_i} \\
&\quad \times \left[\left(\frac{\tilde{\beta}_i}{\tilde{\beta}_i + \left(S_i(t) e^{-\lambda_i(t-s)} + S_i(s)\right) \tilde{\sigma}_{i,i}}\right)^{\tilde{\alpha}_i/\lambda_i}\right. \\
&\quad \times \left(\frac{\tilde{\beta}_{n+1}}{\tilde{\beta}_{n+1} + \left(S_i(t) e^{-\lambda_i(t-s)} + S_i(s)\right) \tilde{\sigma}_{i,n+1}}\right)^{\tilde{\alpha}_{n+1}/\lambda_i} \\
&\quad - \left(\frac{\tilde{\beta}_i^2}{\left(\tilde{\beta}_i + \tilde{\sigma}_{i,i} S_i(t) e^{-\lambda_i(t-s)}\right)\left(\tilde{\beta}_i + \tilde{\sigma}_{i,i} S_i(s)\right)}\right)^{\tilde{\alpha}_i/\lambda_i} \\
&\quad \left. \times \left(\frac{\tilde{\beta}_{n+1}^2}{\left(\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_i(t) e^{-\lambda_i(t-s)}\right)\left(\tilde{\beta}_{n+1} + \tilde{\sigma}_{i,n+1} S_i(s)\right)}\right)^{\tilde{\alpha}_{n+1}/\lambda_i}\right]
\end{aligned} \tag{C.8}$$

for $i = 1, \ldots, n$.

Identification of Parameters in the Gamma Model

Let Λ_{var} be the $n \times n$ matrix given by

$$(\Lambda_{var})_{i,j} = \frac{1}{\lambda_i + \lambda_j}.$$

Furthermore, denote by $\circ$ the Hadamard product. Then the following result will be used to estimate the parameters of the gamma model. Again, we relegate the proof of Proposition 3.3 to the Appendix.

Proposition 3.3. *The mean of X is*

$$\mathbb{E}[X(0)] = \Lambda^{-1}\Sigma_L\beta/2 \tag{C.9}$$

and the auto-covariance of X is

$$\operatorname{cov}(X(0), X(t)) = \left(\Lambda_{var} \circ \left(\Sigma_L\Sigma_L^\top\right)\right) e^{-\Lambda t} \tag{C.10}$$

for $t \geq 0$.

The parameters of the gamma model will be estimated in three steps. First, the mean-reversion matrix Λ will be fitted to the empirical auto-correlation function based on the first 25 lags. From (C.10), it follows that the model auto-correlation function of X_i is $t \mapsto e^{-\lambda_i t}$. To find $\hat{\lambda}_i$, the estimate of λ_i, we therefore minimize

$$\sum_{t=1}^{25} \left(\hat{\rho}_i(t) - e^{-\hat{\lambda}_i t}\right)^2$$

such that $\hat{\lambda}_i > 0$ for $i = 1, \ldots, n$, where $\hat{\rho}_i(t)$ is the empirical auto-correlation function of X_i.

Next, $\hat{\Sigma}_L$ is chosen such that the model matches the empirical covariances. In particular, we choose $\hat{\Sigma}_L$ to minimize

$$\left\|\hat{\Sigma}_X - \Lambda_{var} \circ \left(\hat{\Sigma}_L\hat{\Sigma}_L^\top\right)\right\|^2$$

where $\hat{\Sigma}_X$ is the sample covariance of X, $\|\cdot\|$ is the Frobenius norm and the minimization is done over matrices $\hat{\Sigma}_L$ with non-negative entries of the form in (C.5).

Finally, we discuss how the parameters α and β are estimated. First, we choose $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_n)$ to minimize

$$\left\|\hat{\mu}_X - \Lambda^{-1}\hat{\Sigma}_L\hat{\beta}/2\right\|^2$$

such that $\hat{\beta}_i > 0$, where $\hat{\mu}_X$ is the empirical mean of X.

It is not too difficult to show that $\operatorname{var}(L_i(1)) = 2\alpha_i/\beta_i^2$ and, since the compound Poisson processes are assumed to have unit variance, it therefore follows that

$$1 = \operatorname{var}(L_i(1)) = \frac{2\alpha_i}{\beta_i^2}$$

Consequently, we take $\hat{\alpha}_i = \hat{\beta}_i^2/2$.

3.3 A Lognormal Model

In this section we present a lognormal model relying on different assumptions than the gamma model. We assume that $G(t) := \log X(t)$ can be modelled as a multidimensional Gaussian Ornstein-Uhlenbeck process,

$$dG(t) = -\Upsilon(G(t) - \Theta)dt + \Sigma dB(t), \tag{C.11}$$

where $(B(t))_{t\in\mathbb{R}}$ is an n-dimensional Brownian motion, $\Upsilon \in \mathbb{R}^{n\times n}$ is a diagonal matrix, $\Sigma \in \mathbb{R}^{n\times n}$ is a lower triangular matrix, and $\Theta \in \mathbb{R}^n$.

It is well-known (see e.g. Halliwell (2015)) that the stationary distribution of $G(t)$, when the diagonal elements of Υ all are positive, is normal with mean Θ. The autocovariance of $G(t)$ is given in, for example, Halliwell (2015), and it is the same as for the gamma model found in Prop. 3.3. In particular, we find

$$\Sigma_G(t) := \operatorname{cov}(G(0), G(t))) = \left(\Upsilon_{var} \circ \left(\Sigma\Sigma^\top\right)\right) e^{-\Upsilon t}, \quad t \geq 0. \tag{C.12}$$

Here, Υ_{var} is the $n \times n$ matrix given by

$$(\Upsilon_{var})_{i,j} = \frac{1}{\upsilon_i + \upsilon_j},$$

where υ_i is the i'th entry of Υ, $i = 1, \ldots, n$.

Consequently, the stationary distribution of $X(t)$ is multivariate lognormal with expected value of $X_i(0)$ being

$$\mathbb{E}[X_i(0)] = \exp\left(\Theta_i + \frac{\Sigma_G(0)_{ii}}{2}\right),$$

while the autocovariance is

$$\operatorname{cov}(X_i(0), X_j(t)) = \mathbb{E}[X_i(0)]\mathbb{E}[X_j(0)](e^{\Sigma_G(t)_{ij}} - 1) \tag{C.13}$$

for $i, j = 1, \ldots, n$ (see e.g. Halliwell (2015) for more information on the multivariate lognormal distribution). This implies that the autocorrelation of $X(t)$ is

$$\operatorname{corr}(X_i(0), X_j(t)) = \frac{\exp\left(\Sigma_G(0)_{ij}e^{-t\upsilon_j}\right) - 1}{\sqrt{(e^{\Sigma_G(0)_{ii}} - 1)(e^{\Sigma_G(0)_{jj}} - 1)}}, \tag{C.14}$$

for $i, j = 1, \ldots, n$.

Distribution of $P_i(t)$ in the Lognormal Model

Having the results for $X(t)$ from the previous section in mind, the stationary distribution of $P_i(t)$ follows and is given in Prop. 3.4.

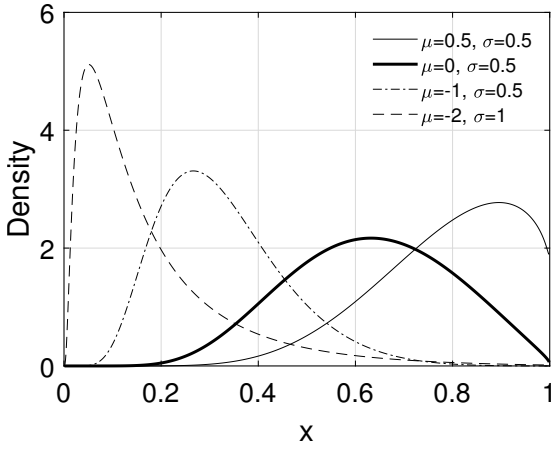


Fig. C.4: Different variations of the density in Eq. (C.16).

Proposition 3.4. *The stationary distribution of $P_i(t)$ is characterized by the density*

$$f_{P_i(t)}(x) = \frac{-1}{(1-x)\log(1-x)\sqrt{\Sigma_G(0)_{i,i}}}\phi\left(\frac{\log\left(-\frac{\log(1-x)}{S_i(t)}\right)-\Theta_i}{\sqrt{\Sigma_G(0)_{i,i}}}\right), \quad \text{(C.15)}$$

where $\Sigma_G(0)_{i,i}$ is the i'th element of the diagonal of $\Sigma_G(0)$, Θ_i is the i'th element of Θ, and $\phi(\cdot)$ is the density of the standard normal distribution.

To investigate the density of $P_i(t)$ in more detail, consider for a moment a more generic version of Eq. (C.15), given by

$$f(x|\mu,\sigma) = \frac{-1}{(1-x)\log(1-x)\sigma}\phi\left(\frac{\log\left(-\log(1-x)\right)-\mu}{\sigma}\right). \quad \text{(C.16)}$$

As can be seen in Fig. C.4, showing examples of the density given different values of μ and σ in Eq. (C.16), the distribution is rather flexible and capable of attaining quite different forms.

Covariance Between Wind Indexes in the Lognormal Model

Deriving the covariances between wind indexes in the lognormal model is closely related to the derivation of the Laplace transform of the lognormal distribution. To the best of our knowledge, no closed-form has been derived for the Laplace transform of the lognormal distribution, but there exist approximations, see e.g. Asmussen *et al.* (2016). With regard to this paper, we refer the interested reader to Asmussen *et al.* (2016) and the references herein for further information, and employ numerical integration by exploiting our knowledge of the distribution of $G(t)$ to determine the covariances between the wind indexes.

Identification of Parameters in the Lognormal Model

To identify the parameters of the model, we employ the method of moments as in the gamma model case. We first identify $\Sigma_G(0)$ by exploiting Eq. (C.13),

$$\Sigma_G(0)_{ij} = \log\left(\frac{\hat{\Sigma}_{X,ij}}{\hat{\mu}_i\hat{\mu}_j} + 1\right), \tag{C.17}$$

with $\hat{\mu}_i$ being the empirical mean of $X_i(t)$, and $\hat{\Sigma}_{X,ij}$ is the (i,j)th entry of the empirical covariance between $X_i(0)$ and $X_j(0)$.

Having obtained an estimate of $\Sigma_G(0)$ and remembering the model implied autocorrelation in Eq. (C.14), we identify v_i by minimizing

$$\sum_{t=1}^{25}\left(\hat{\rho}_i(t) - \frac{\exp\left(\hat{\Sigma}_G(0)_{ii}e^{-tv_i}\right) - 1}{\sqrt{(e^{\hat{\Sigma}_G(0)_{ii}} - 1)(e^{\hat{\Sigma}_G(0)_{ii}} - 1)}}\right)^2,$$

where $\hat{\rho}_i(t)$ is the empirical autocorrelation function of $X_i(0)$ and $X_i(t)$. Here, as in the gamma model, we use the first 25 lags of the empirical autocovariance function to estimate λ_i. With $\hat{\Upsilon}$, consisting of the estimated v_i for $i = 1,\ldots,n$ in the diagonal, at hand, we identify $\Sigma\Sigma^\top$ by

$$\Sigma\Sigma^\top = \hat{\Sigma}_G(0) \oslash \hat{\Upsilon}_{var},$$

where $\oslash$ is the Hadamard division defined for two matrices A and B by $A \oslash B = A_{ij}/B_{ij}$. Lastly, we determine Θ by

$$\Theta_i = \log(\mu_i) - \frac{\hat{\Sigma}_G(0)_{ii}}{2}, \quad i = 1,\ldots,n. \tag{C.18}$$

3.4 Comparison of the Gamma and Lognormal Model

The covariances between indexes in the gamma model can be calculated fast using Proposition 3.2 to find the optimal hedging strategy (see Sec. 5). The noise in the gamma model also has a compelling interpretation, where an idiosyncratic risk is associated to each site index and a systematic risk is associated to all site indexes and the German index. On the other hand, the lognormal model gives rise to closed-form expressions of the densities of all indexes as opposed to only the German index in the gamma model. The lognormal model is simple in the sense that the underlying process is a Gaussian driven OU process. This makes it possible to do numerical analysis based on Gaussian theory.

Both the gamma and lognormal model have straightforward and fast estimation procedures, making them easy to implement. Furthermore, as we will see in Sec. 5.1, both models capture the autocorrelation of X_i, the cross-autocorrelations between X_i and X_j, and the stationary distribution of X_i well.

	$\hat{a}_i$	$\hat{b}_i$	$\hat{c}_i$
Site 1	0.1721	-0.0491	-0.0804
Site 2	0.2848	-0.0405	-0.0956
Site 3	0.2294	-0.0322	-0.1226
German	0.2732	-0.0298	-0.1285

Table C.1: Fitted seasonal parameters for the four wind indexes.

	$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\lambda}_i$	$\hat{\sigma}_{i,i}$	$\hat{\sigma}_{i,4}$
Site 1	0.0271	0.2328	0.8977	1.0305	1.1593
Site 2	0.0538	0.3282	0.7589	0.6101	0.9792
Site 3	0.1383	0.5260	0.8513	1.1674	0.8247
German	0.8960	1.3387	0.6539	(-)	0.9781

Table C.2: Estimated parameters in the gamma model.

4 Estimation Results

In this section we summarize and discuss the estimation results. As a starting point, we consider the fitted seasonal functions. In Table C.1 we report the fitted parameters for all four wind indexes.

4.1 Gamma Model

Fig. C.5 shows the theoretical autocorrelation implied by the estimated gamma model compared to the empirical autocorrelation. The fit to the empirical autocorrelation is convincing, and it is worth noticing that the cross-autocorrelations match well even though the model has only been estimated to the marginal autocorrelation functions.

In Fig. C.6 the histogram of X_i and the model density based on a simulation are shown. We see that the distribution of the data is captured well by the model.

We report the estimated parameters in Table C.2. The parameters α_4 and β_4 are considerably larger than α_i and β_i for $i = 1, 2, 3$. This implies that the systematic risk factor L_4 jumps much more frequently than L_i, $i = 1, 2, 3$, but that the jumps of L_4 are relatively small compared to the jumps of L_i, $i = 1, 2, 3$. This aligns well with the intuition that the systematic risk is associated to the wind utilization of the whole of Germany.

To further assess the model, we report in Table C.3 the mean, variance, skewness, and kurtosis of the gamma model along with the empirical and lognormal equivalents for the German wind index[2]. The gamma model cap-

[2]Since the same quantities for the site wind indexes are not relevant in the remaining part of the paper, we have chosen to omit them.

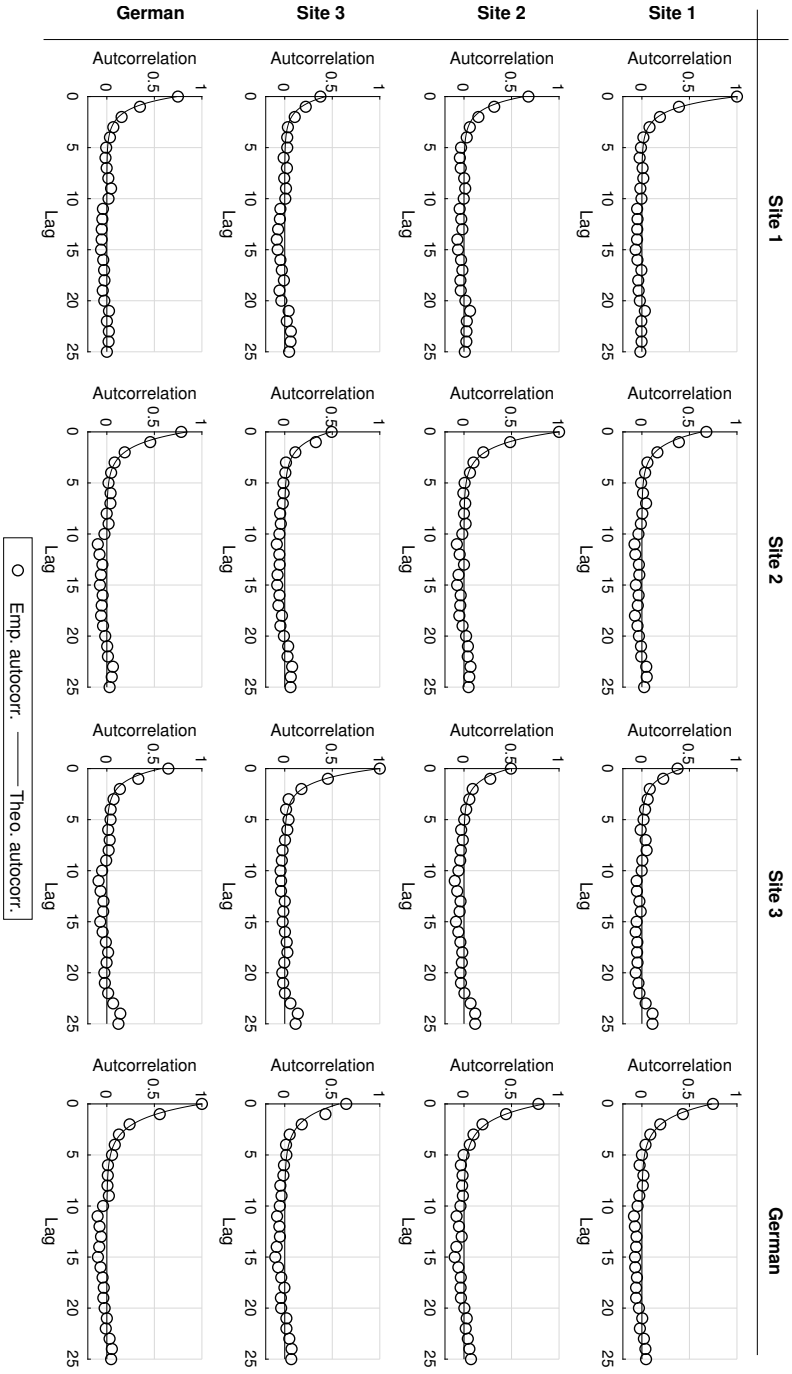
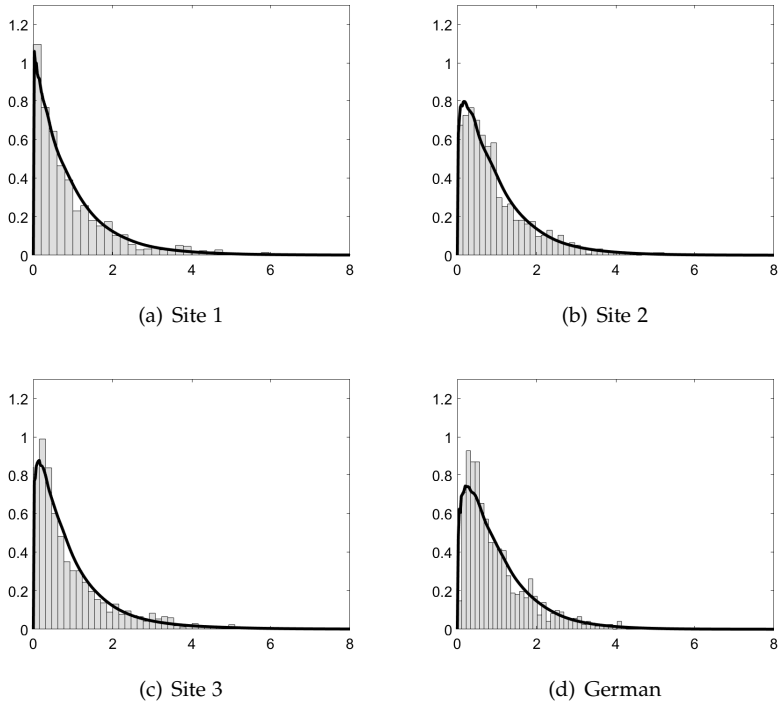


Fig. C.5: Empirical autocorrelation and theoretical autocorrelation implied by the fitted gamma model. The (i, j)'th panel shows $\text{cor}(X_i(0), X_j(t))$ for $t = 0, 1, \ldots, 25$.



(a) Site 1 (b) Site 2

(c) Site 3 (d) German

Fig. C.6: Histograms of $X_i(t)$ with the fitted densities of the gamma model.

	Mean	Variance	Skewness	Kurtosis
Gamma	1.00	0.73	1.71	7.38
Lognormal	1.00	0.73	3.17	24.98
Empiricial	1.00	0.73	1.67	6.27

Table C.3: Mean, variance, skewness, and kurtosis of the German wind index in the gamma and lognormal model together with the empirical values for the German wind index.

	$\hat{\Theta}_i$	$\hat{\upsilon}_i$
Site 1	-0.4282	0.7080
Site 2	-0.3215	0.6341
Site 3	-0.3803	0.6837
German	-0.2711	0.5607

Table C.4: Estimated parameters in the lognormal model.

tures the first two moments very well as expected from the estimation procedure, where we match the gamma model to the empirical mean and variance. Further, the empirical skewness and kurtosis are also captured very well by the gamma model.

4.2 Lognormal Model

Fig. C.7 shows the theoretical autocorrelation implied by the estimated lognormal model compared to the empirical autocorrelation. As in the gamma model case, the lognormal model captures the autocorrelation and cross-autocorrelation well, in particular taking into account that only the autocorrelation is used to estimate the parameters affecting both the autocorrelations and the cross-autocorrelations.

Fig. C.8 shows histograms of the marginal distributions and the fitted lognormal densities. The lognormal model provides overall a decent fit, but seems to capture the distribution of the German wind index better than the site indexes. The estimated Θ and Υ for the lognormal model is reported in Table C.4 and the estimated Σ is

$$\hat{\Sigma} = \begin{bmatrix} 1.0987 & 0 & 0 & 0 \\ 0.6763 & 0.5886 & 0 & 0 \\ 0.4902 & 0.3505 & 0.8376 & 0 \\ 0.6381 & 0.2539 & 0.2162 & 0.3035 \end{bmatrix}.$$

Although the speed of mean reversion parameters $\hat{\upsilon}_i$ differ in the lognormal model compared to the gamma model, the same pattern is observed, with the German wind index being the most persistent.

Returning to Table C.3, the lognormal model matches the empirical mean and variance as a results of the estimation procedure, but it does not capture

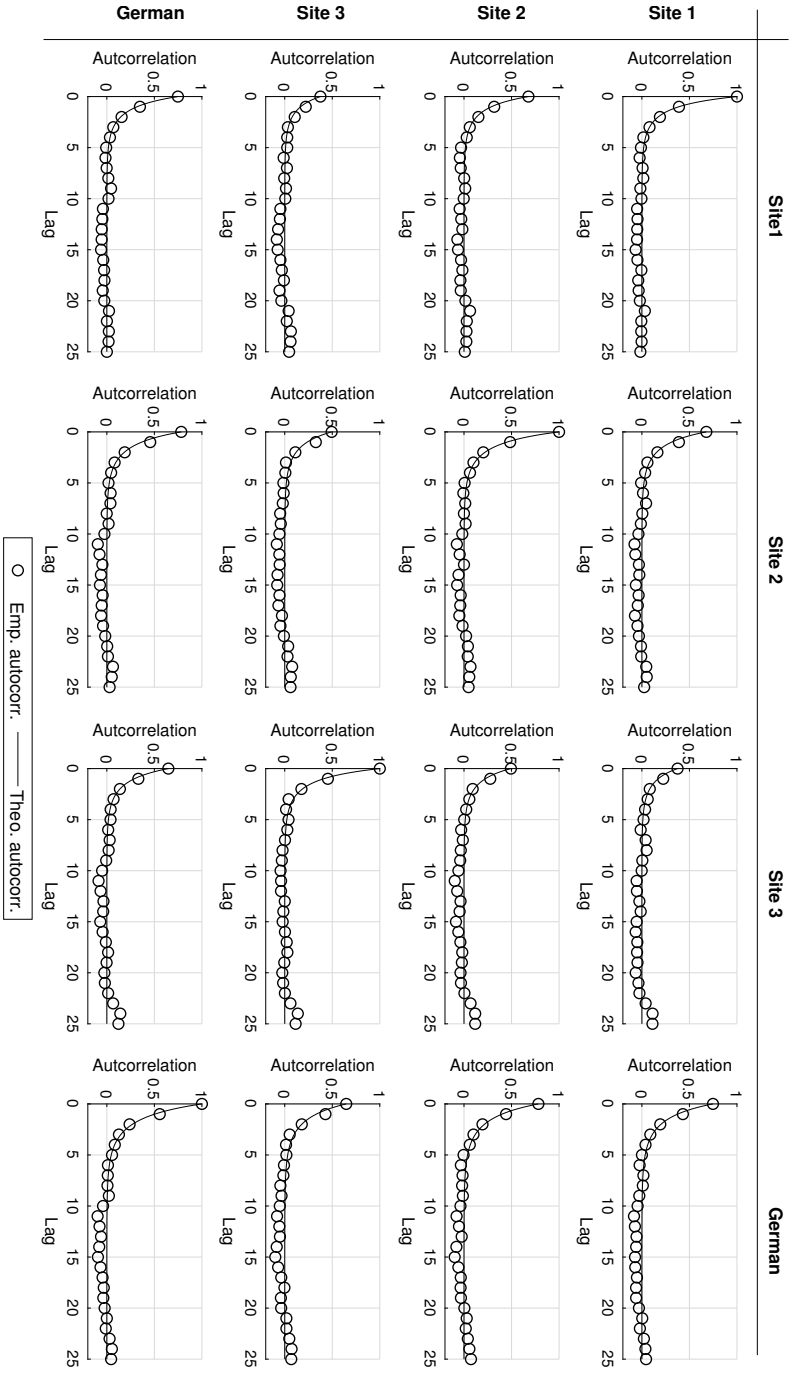
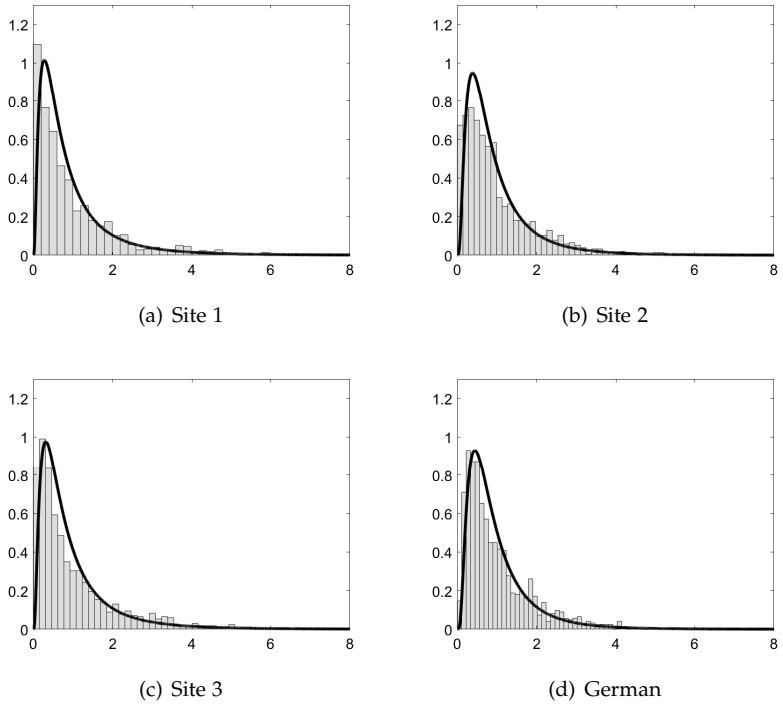


Fig. C.7: Empirical autocorrelation and theoretical autocorrelation implied by the fitted lognormal model. The (i, j)'th panel shows $\text{cor}(X_i(0), X_j(t))$ for $t = 0, 1, \ldots, 25$.



(a) Site 1 (b) Site 2

(c) Site 3 (d) German

Fig. C.8: Histograms of $X_i(t)$ with fitted lognormal densities.

the higher order standardized moments. This indicates that the lognormal model does not capture the whole distribution of the data as well as the gamma model.

5 Hedging Wind Power Production

In the following we denote the German wind index at day t by $P_n(t)$. An exchange-traded WPF contract is written on the underlying daily wind index, $P_n(t)$. The payoff of a long position in such a contract is

$$H(\bar{P}_n(S,T) - P_n(t_0,S,T))X, \tag{C.19}$$

where H is the number of hours during the delivery period $[S,T]$, $P_n(t_0,S,T)$ is the WPF price agreed on at time $t_0 < S < T$, X is a specified tick size, and

$$\bar{P}_n(S,T) = \frac{1}{T-S+1}\sum_{t=S}^{T} P_n(t).$$

From Eq. (C.19) it is apparent that a short position results in a positive payoff in low wind scenarios according to the short position equivalent to Eq. (C.19),

$$H(P_n(t_0,S,T) - \bar{P}_n(S,T))X.$$

That is, if the realization of $\bar{P}_n(S,T)$ is lower than $P_n(t_0,S,T)$. This is favourable for a WPP, since this payoff will offset the loss in income from the long position in wind power production.

To be more specific, let C_i denote the capacity of WPP i, and let $P_i(t)$ denote the daily wind index/utilization of WPP i such that $C_iP_i(t)$ is the actual production of power. Further assume that the WPP receives a fixed price Q_i per produced MWh. The long position in wind power production for WPP i doing the period $[S,T]$ is therefore

$$\bar{P}_i(S,T)C_iHQ_i, \tag{C.20}$$

where

$$\bar{P}_i(S,T) = \frac{1}{T-S+1}\sum_{t=S}^{T} P_i(t). \tag{C.21}$$

Assume that the WPP takes a position $\gamma_i \in \mathbb{Z}$ in WPF contracts with delivery period being $[S,T]$. The payoff from taking this position and the long position in wind power production results in a portfolio with payoff

$$H\bar{P}_i(S,T)C_iQ_i + \gamma_i H(\bar{P}_n(S,T) - P_n(t_0,S,T))X. \tag{C.22}$$

From Eq. (C.22) it is clear that perfectly hedging the volumetric risk would mean to choose γ_i such that $H\bar{P}_i(S,T)C_iQ_i = -\gamma_i H\bar{P}_n(S,T)X$, resulting in the deterministic payoff $\gamma_i HP_n(t_0,S,T)X$. However, the problem for the WPP is that the stochastic terms $\bar{P}_i(S,T)$ and $\bar{P}_n(S,T)$ are not perfectly dependent, and hence obtaining the deterministic payoff $\gamma_i HP_n(t_0,S,T)X$ is not possible. In fact, as shown in Christensen & Pircalabu (2018), it is far from optimal using the exchange-traded WPF contracts for hedging purposes for a single WPP, depending on the dependence structure between the site-specific wind index and the underlying index of the WPF contract.

5.1 Perfect Hedging of Volumetric Risk Using Tailor-Made Wind Power Futures

Tailor-made over-the-counter WPF contracts is a way of perfectly hedging the volumetric risk. Instead of going short the exchange-traded WPF contract, the WPP could instead go short an over-the-counter WPF contract with the underlying being $P_i(t)$ instead of $P_n(t)$. In the following we therefore consider the situation of an energy management company (EMC) acting as counterparty of these tailor-made WPF contracts from $n-1$ different WPPs in Germany. Let $H(\bar{P}_i(S,T) - P_i(t_0,S,T))C_iQ_i$ be the payoff of a long position in a tailor-made WPF contract for WPP i. Thus, from the point of view of the EMC, the payoff of acting as counterparty for $n-1$ different WPPs and taking a position $\gamma \in \mathbb{Z}$ in the exchange-traded WPF contract is

$$R_C(\gamma) = \sum_{i=1}^{n-1} H(\bar{P}_i(S,T) - P_i(t_0,S,T))C_iQ_i + \gamma H(\bar{P}_n(S,T) - P_n(t_0,S,T))X, \tag{C.23}$$

while the payoff from the point of view of the ith WPP is

$$\begin{aligned} R_{WPP,i} &= H\bar{P}_i(S,T)C_iQ_i + H(P_i(t_0,S,T) - \bar{P}_i(S,T))C_iQ_i \\ &= HP_i(t_0,S,T)C_iQ_i. \end{aligned}$$

We argue that this construction can be beneficial for both the individual WPPs and the EMC: Firstly, the individual WPPs obtain a perfect hedge of their volumetric risk, and secondly, with an appropriate number of WPPs and distribution of the WPPs geographically, the portfolio of tailor-made WPF contracts approximately replicates the exchange-traded WPF contract.

5.2 Minimum Variance Hedge of a Tailor-Made WPF Contracts Portfolio

In this section we discuss a minimum variance hedge of a portfolio consisting of tailor-made WPF contracts for the EMC. I.e., from Eq. (C.23) we define the

objective to

$$\min_{\gamma} \operatorname{var}(R_C(\gamma)).$$

The variance is

$$\begin{aligned}
\operatorname{var}(R_C(\gamma)) = \operatorname{var}\Bigg[& \sum_{i=1}^{n-1} H\left(\frac{1}{T-S+1}\sum_{t=S}^{T} P_i(t) - P_i(t_0,S,T)\right) C_i Q_i \\
& + \gamma H\left(\frac{1}{T-S+1}\sum_{t=S}^{T} P_n(t) - P_n(t_0,S,T)\right) X\Bigg] \\
= & \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(\frac{H}{T-S+1}\right)^2 C_i Q_i C_j Q_j \sum_{t=S}^{T}\sum_{s=S}^{T} \operatorname{cov}(P_i(t), P_j(s)) \\
& + \left(\gamma \frac{H}{T-S+1} X\right)^2 \sum_{t=S}^{T}\sum_{s=S}^{T} \operatorname{cov}(P_n(t), P_n(s)) \\
& + 2\sum_{i=1}^{n-1}\left(\frac{H}{T-S+1}\right)^2 \gamma X C_i Q_i \sum_{t=S}^{T}\sum_{s=S}^{T} \operatorname{cov}(P_n(t), P_i(s)). \quad \text{(C.24)}
\end{aligned}$$

It follows from Eq. (C.24) that the optimal position of WPF contracts is

$$\gamma = -\frac{\sum_{i=1}^{n-1} C_i Q_i \sum_{t=S}^{T}\sum_{s=S}^{T} \operatorname{cov}(P_n(t), P_i(s))}{X \sum_{t=S}^{T}\sum_{s=S}^{T} \operatorname{cov}(P_n(t), P_n(s))}. \quad \text{(C.25)}$$

Besides the fact that the dependencies between the stochastic variables impact the size of γ, the size of each wind site measured by C_i and the price paid for each MWh to each wind site measured by Q_i both translate linearly to the size of γ. Therefore, the larger the wind site or the higher the price paid for each MWh, the larger γ will be in absolute terms (all other things being equal).

In-Sample Hedging Effectiveness

In the following we consider the case of an EMC that needs to hedge its portfolio of tailor-made WPF from one year ahead to two years ahead. The considered wind sites are the ones depicted in Fig. B.1(a). We assume that the contract specifications for each site is as shown in Table C.5. Further, we assume that $X = 100$ EUR. The estimated parameters of the gamma and lognormal model are the ones reported in Sec. 5.1.

In Table C.6 we present the hedging results for the gamma and lognormal model. We include the case with all three sites and the German WPF in the portfolio, and then three cases where we only include one of the wind sites and the German WPF. In each case, we report the

Site ID	Capacity in MW, C_i	Price in EUR/MWh, Q_i
1	100	30
2	100	30
3	100	30

Table C.5: Fictional contract specifications for the sites in Fig. B.1(a).

Case	Sites in portfolio	$\hat{\gamma}$	$\text{var}(R_C(0))$	$\text{var}(R_C(\hat{\gamma}))$	Variance reduction (%)
			Gamma		
1	1,2,3	-63.60	$8.12 \cdot 10^{11}$	$1.31 \cdot 10^{11}$	83.83
2	1	-18.87	$8.55 \cdot 10^{10}$	$2.56 \cdot 10^{10}$	70.12
3	2	-26.79	$1.55 \cdot 10^{11}$	$3.41 \cdot 10^{10}$	77.96
4	3	-17.94	$1.25 \cdot 10^{11}$	$7.07 \cdot 10^{10}$	43.42
			Lognormal		
1	1,2,3	-64.00	$7.06 \cdot 10^{11}$	$1.21 \cdot 10^{11}$	82.89
2	1	-18.97	$8.12 \cdot 10^{10}$	$2.98 \cdot 10^{10}$	63.33
3	2	-25.10	$1.34 \cdot 10^{11}$	$4.42 \cdot 10^{10}$	67.04
4	3	-19.92	$1.13 \cdot 10^{11}$	$5.66 \cdot 10^{10}$	50.05

Table C.6: Optimal hedging quantity γ implied by the gamma and lognormal model for different portfolios consisting of different wind sites, and the corresponding variance of the portfolio excluding the exchange-traded WPF contract, and the variance of the portfolio when the optimal hedge is employed. Additionally, we show in all cases the associated variance reduction in percentage.

model-implied optimal position of exchange-traded WPF contracts, $\hat{\gamma}$, and the variance reduction (in percentage) implied by the model calculated by $[\text{var}(R_C(0)) - \text{var}(R_C(\hat{\gamma}))]/\text{var}(R_C(0))$. It is apparent that the portfolio with all three sites outperforms the three other cases, confirming the diversification approach of the EMC discussed in Sec. 5.1.

The fact that the difference regarding $\hat{\gamma}$ is small indicate that both models could be used interchangeably to determine an appropriate hedge, though the difference in variance reduction will mislead in a risk management context. In other words, if the wind indexes are driven by the gamma (lognormal) model, and one uses the lognormal (gamma) model to determine hedges, the variance reduction implied by the used model is wrong, while the hedging quantity is relatively close to the optimal hedge.

Comparing Eq. (C.22) to Eq. (C.23), the cases 2, 3, and 4 represent the variance reduction implied by the model if the individual wind sites were to hedge their power production themselves by using the exchange-traded WPF contract. From a social welfare point of view, the sum of variances of case 2, 3, and 4 is approximately 8% larger for both models. So not only does the model suggest that tailor-made WPF contracts constitute an obvious way of mitigating uncertainty for wind power producers, but also as a way of

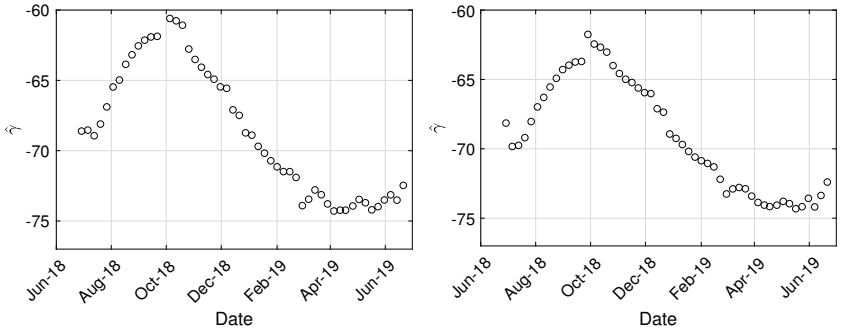
optimizing the integration of wind power penetration in the electricity grid.

Out-Of-Sample Hedging Effectiveness

In this section we consider the same portfolio of wind sites as in the previous case study, and the specifications of the sites are therefore specified in Table C.5. However, here we assess the model on out-of-sample observations. We assume that an EMC has bought tailor-made WPF contracts at the three sites for the period from 2 July 2018 to 30 June 2019, corresponding to 364 days or 52 weeks. We employ a weekly minimum variance hedging strategy, meaning that the EMC has a naked position in a portfolio of tailor-made WPF contracts for the entire period with the exception of the front week. To concretize, the first position taken in exchange-traded WPF contracts is the contract with a weekly delivery period from 2 July 2018 to 8 July 2018. The position is taken based on a model that is estimated by using two years of observations ending the last trading day before the delivery period of the weekly exchange-traded WPF contract. With the delivery period starting the 2 July 2018, the last trading day turns out to be 29 June 2018. Then we step one week ahead and determine the appropriate hedge for the week starting 9 July 2018 and ending 15 July 2018, but again only by employing two years of in-sample observations to estimate the model (the estimation period again ends on the last day where one can trade the weekly exchange-traded WPF contract). In this way we end up with 52 hedging quantities, where each quantity is calculated using different estimated parameters of the model due to the moving two-year observation period.

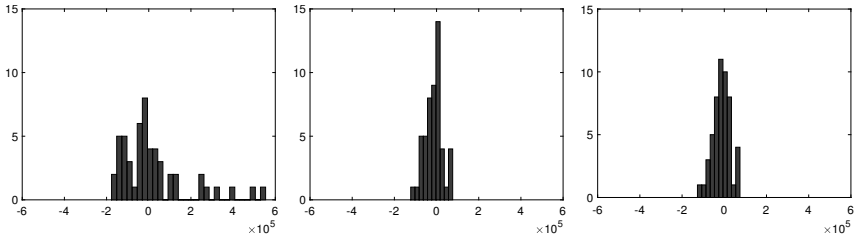
A comment on the model specifications is in place. The stationarity of the models in Sec. 3 might seem unreasonable in the present context, given the short period of time between an estimation date and the corresponding start date of delivery of the exchange-traded WPF contract. However, we also implemented the models that take the conditional distribution into account, resulting in similar results. For the sake of keeping the presentation as clear as possible, we have therefore only chosen to present the stationary versions of the models.

The resulting optimal hedge quantities are depicted in Fig. C.9, indicating a seasonal pattern with more exchange-traded WPF contracts needed during spring compared to autumn. Considering Eq. (C.25), this is the result of the fact that the difference between the sum of the autocovariances of the German wind index and the sum of the autocovariances between the German wind index and the site indexes increases. To assess the hedging effectiveness, we calculate the corresponding implied weekly payoff, $R_C(\hat{\gamma})$, for each weekly hedge quantity, $\hat{\gamma}$. Since we have a variance minimizing perspective, we force a simplistic view on $P_i(t_0, S, T)$ for all wind indexes. Specifically, we assume that for each i, $P_i(t_0, S, T)$ for all weeks during the out-of-sample period from



(a) Gamma model (b) Lognormal model

Fig. C.9: Variance minimizing hedge quantitiy, $\hat{\gamma}$, implied by (a) the gamma model and (b) the lognormal model for the 52 weeks covering the period from 2 July 2018 to 30 June 2019.



(a) No hedge (b) With hedge, gamma model (c) With hedge, lognormal model

Fig. C.10: Histograms of (a) the payoff for EMC by not hedging the portfolio of tailor-made WPF contracts with exchange-traded WPF contracts, and (b)-(c) using the gamma and lognormal model to find the position in exchange-traded WPF contracts used as a hedging instrument for the portfolio of tailor-made WPF contracts.

2 July 2018 to 30 June 2019 is the mean of $P_i(t)$ over the first estimation period spanning 1 July 2016 to 29 June 2018,

Fig. C.10 shows a histogram of the payoffs of the portfolio of tailor-made WPF contracts and the exchange-traded WPF contract acting as hedging instrument. Compared to Fig. C.10(a), the variances in Fig. C.10(b) and Fig. C.10(c) are clearly reduced. In fact, the variance reduction in percentage of using the exchange-traded WPF contracts as hedging instrument is 93.64% for the gamma model and 93.62% for the lognormal model.

Risk Premium of Wind Power Futures

Since tailor-made WPF contracts are, by construction, traded over-the-counter, it is worth to discuss the risk premium of such contracts. As a refer-

ence point, we consider the risk premium of the exchange-traded WPF contracts. We define the risk premium as the model implied WPF contract price under the physical measure subtracted from the observed exchange-traded WPF contract price. The model implied price is defined by $\mathbb{E}[\bar{P}_n(S,T)]$, meaning that the risk premium $RP(t_o,S,T)$ is

$$RP(t_0,S,T) = \bar{P}_n(t_0,S,T) - \mathbb{E}[\bar{P}_n(S,T)] \tag{C.26}$$

on day t_0 for the delivery period $[S,T]$. The observed quoted exchange-traded WPF prices are obtained from NASDAQ OMX. As concluded in Sec. 5.2, the stationarity of the models does not imply different results compared to the conditional versions of the models for such long time periods, so to ease the presentation, we only consider the unconditional expected value here[3].

We limit ourselves to yearly and quarterly exchange-traded WPF contracts for two reasons. First, it is unlikely that the tailor-made WPF contracts in general will be specified for a short delivery such as a week as a result of such non-standardized instrument. Secondly, as concluded in Bertsekas & Pircalabu (2018), fundamentals impact the information premium of exchange-traded WPF contracts with a short delivery period (e.g. a week) and a short period of time to delivery, which we would like to avoid. Thirdly, to assess the seasonal differences we also consider quarterly contracts.

For the period from 1 July 2016 to 30 June 2019, we show $\mathbb{E}[\bar{P}_n(S,T)]$ and $\bar{P}_n(t_0,S,T)$ in Fig. C.11 for the front year (that is, for a given date, the front year denotes the next year). The quoted prices are fairly constant throughout the entire period, which could be a consequence of illiquidity of exchange-traded WPF contracts. The risk premium is -0.011 for the gamma model and -0.013 for the lognormal model on average. Since we are considering a yearly WPF contract we can ignore the seasonality and use the empirical mean to assess the risk premium. The empirical risk premium is -0.011, agreeing with the gamma model. This is likely a consequence of the gamma model having a better fit to the distribution of the German wind index as discussed in Sec. 5.1 (see also Table C.3).

Fig. C.12(a) shows the model implied and quoted prices for the front quarter, and Fig. C.12(b) shows the corresponding risk premium. The mean of the risk premium in this case is -0.014 for the gamma model and -0.016 for the lognormal model. The seasonal variation in the prices peaks for contracts with delivery during Q4 and Q1, simply since more wind is present during these quarters. This is also reflected in the model-implied prices. The peaks in the risk premium are observed for contracts with delivery during Q1 and Q2. One explanation of this could be non-aligned incentives to engage in the

[3]Despite the fact that $RP(t_0,S,T)$ still depends on t_0 thorugh $\bar{P}_n(t_0,S,T)$, the assumption of stationarity is to some degree confirmed by the constant pattern of $\bar{P}_n(t_0,S,T)$ observed in Figs. C.11 and C.12(a).

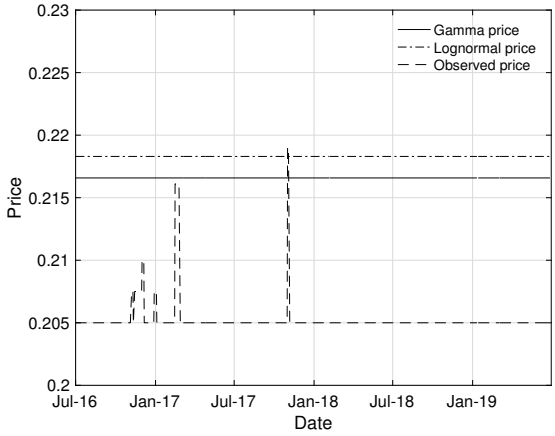


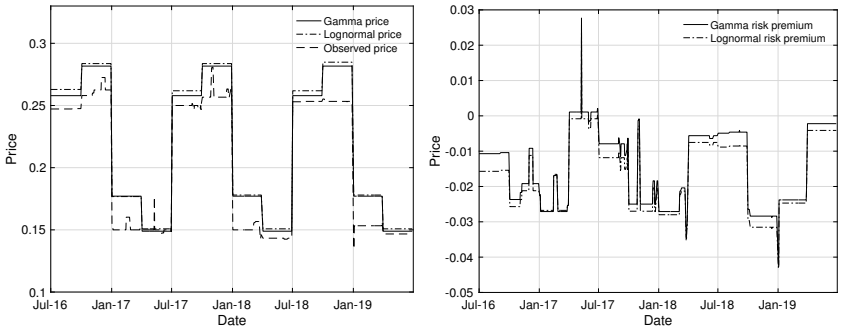
Fig. C.11: The model implied and quoted price for the front year for the period from 1 July 2016 to 30 June 2019. Notice that the date refers to the observation date; i.e., the date where the contract is quoted.

WPF market throughout the year for the buying and selling side. Christensen *et al.* (2019) shows that the hedging benefits are greater for CPPs during Q3 and Q4 compared to Q1 and Q2; hence, during Q3 and Q4, CPPs are more interested in WPF contracts and thus willing to pay more.

A negative risk premium is in line with the findings in Bertsekas & Pircalabu (2018) and Gersema & Wozabal (2017). One might argue that this is expected from a hedging benefit perspective, since the hedging benefits in general are greater for the selling side than the buying side (see Christensen & Pircalabu (2018), Christensen *et al.* (2019), and Gersema & Wozabal (2017)). Continuing this argument, the risk premium is likely to be even more negative in the tailor-made WPF contracts market as a result of the perfect hedge implied by the tailor-made WPF contracts for WPPs. However, from the perspective of the individual WPP, this extra risk premium associated with the tailor-made WPF contract compared to the exchange-traded WPF contract has to be weighted against the deterministic payoff implied by the tailor-made WPF contract.

6 Conclusion

In this paper, we propose and compare two multivariate continuous-time models, the gamma and lognormal model, for the joint behaviour of wind indexes. We discuss the properties of the models, and propose estimation procedures. Empirically, we employ the models to a joint model for the wind indexes at three different wind sites in Germany, and the German wind index that represents the overall utilization of wind power production in Germany.



(a) Model and quoted price (b) Risk premium

Fig. C.12: (a) the model implied and quoted price for the front quarter, and (b) and the risk premium for the front quarter. The observations period is from 1 July 2016 to 30 June 2019. Notice that the date refers to the observation date; i.e., the date where the contract is quoted.

We find that both models are able to capture the autocorrelation structure well. However, the gamma model captures the skewness and kurtosis of the German wind index better than the lognormal model.

The models are applied to a variance-minimizing hedging strategy of a portfolio consisting of long positions in so-called tailor-made wind power futures contracts at the three wind sites, and a short position in the exchange-traded wind power futures contract. The hedging effectiveness is assessed in an in-sample and out-of-sample context. Both models indicate that a significant variance reduction can be obtained by hedging the portfolio with the exchange-traded wind power futures contracts in-sample as well as out-of-sample. Further, the hedging benefits are greater for the portfolio of tailor-made wind power futures compared to hedging each individual wind site with exchange-traded wind power futures contracts.

The risk premium of the exchange-traded wind power futures contracts is examined, where we find that the gamma model implies a more reliable estimate of the risk premium. A negative risk premium is observed in line with other findings in the literature for both yearly and quarterly contracts. Even though the tailor-made wind power futures contracts give each wind power producer a perfect volumetric hedge of her wind power production, we argue that it is likely that the risk premium for a tailor-made wind power futures contract is even more negative compared to the exchange-traded contract.

Acknowledgements

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B. Victor Rohde is supported by the Danish Council

for Independent Research under Grant DFF-4002-00003.

A Theoretical Results for the Gamma Model

This appendix is dedicated to proving Prop. 3.2 and Proposition 3.3. We start by proving Prop. 3.3, which the lemma below is a first step towards. We will use some standard results about continuous-time moving averages, all of which can be found in Rajput & Rosinski (1989).

The following Lemma is well-known, but we give a proof for the sake of completeness.

Lemma A.1. *Let $t \geq 0$ and consider the two one-dimensional processes*

$$Y_1(t) = \int_{-\infty}^{t} f_1(t-u)dZ(u) \quad \textit{and} \quad Y_2(t) = \int_{-\infty}^{t} f_2(t-u)dZ(u) \tag{C.27}$$

for functions f_1 and f_2 in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and where Z is a one-dimensional Lévy process with second moment. Then

$$\mathbb{E}[Y_1(0)] = \int_0^\infty f_1(u)du\mathbb{E}[Z(1)]$$

and

$$\mathbb{E}[(Y_1(0) - \mathbb{E}[Y_1(0)])(Y_2(t) - \mathbb{E}[Y_2(t)])] = \int_0^\infty f_1(u)f_2(t+u)du \operatorname{var}(Z(1)).$$

Proof. Let $\psi_{Y_1(0),Y_2(t)}$ be the cumulant generating function of $(Y_1(0), Y_2(t))$ and ψ_Z be the cumulant generating function of Z. Then

$$\begin{aligned}\psi_{Y_1(0),Y_2(t)}(x) &= \log \mathbb{E}[\exp\{x_1Y_1(0) + x_2Y_2(t)\}] \\ &= \int_0^t \psi_Z(x_2f_2(u))du + \int_0^\infty \psi_Z(x_1f_1(u) + x_2f_2(t+u))du.\end{aligned}$$

It follows that for $n_1, n_2 \in \mathbb{N}_0$ with $n_1 + n_2 \leq 2$,

$$\begin{aligned}\frac{d^{n_1+n_2}}{dx_1^{n_1}dx_2^{n_2}}\psi_{Y_1,Y_2}(x) &= \int_0^t f_2^{n_2}(u)\psi_Z^{(n_2)}(x_2f_2(u))du \\ &\quad + \int_0^\infty f_1^{n_1}(u)f_2^{n_2}(u)\psi_Z^{(n_1+n_2)}(x_1f_1(u) + x_2f_2(u))du.\end{aligned}$$

where $\psi_Z^{(n_1+n_2)}$ denotes the $n_1 + n_2$ times derivative of ψ_Z. We conclude that

$$\mathbb{E}[Y_1(0)] = \frac{d}{dx_1}\psi_{Y_1(0),Y_2(t)}(0) = \int_0^\infty f_1(u)du\mathbb{E}[Z(1)].$$

Assume now, without loss of generality, $\mathbb{E}[Z(1)] = 0$. Then

$$\begin{aligned}
&\mathbb{E}[(Y_1(0) - \mathbb{E}[Y_1(0)])(Y_2(t) - \mathbb{E}[Y_2(t)])] \\
&\quad = \frac{d^2}{dx_1 dx_2} \psi_{Y_1(0), Y_2(t)}(0) \\
&\quad = \int_0^\infty f_1(u) f_2(t+u) du \operatorname{var}[Z(1)]
\end{aligned}$$

□

Proof of Prop. 3.3. Let $\sigma_{i,k}$ denote the (i,k)'th entry of Σ_L. Then, using Lemma A.1,

$$\begin{aligned}
\mathbb{E}[X_i(t)] &= \sum_{k=1}^n \mathbb{E}\left[\int_{-\infty}^t e^{-\lambda_i(t-u)} \sigma_{i,k} dL_k(u)\right] \\
&= \sum_{k=1}^n \frac{1}{\lambda_i} \sigma_{i,k} \beta_k / 2 \\
&= \left(\Lambda^{-1} \Sigma_L \beta / 2\right)_i .
\end{aligned}$$

This gives (C.9). Assume now, without loss of generality, $\mathbb{E}[L(1)] = 0$. Then, using Lemma A.1 again,

$$\begin{aligned}
&\mathbb{E}[X_i(0) X_j(t)] \\
&\quad = \mathbb{E}\left[\left(\sum_{k=1}^n \int_{-\infty}^0 e^{-\lambda_i(-u)} \sigma_{i,k} dL_k(u)\right)\left(\sum_{k=1}^n \int_{-\infty}^t e^{-\lambda_j(t-u)} \sigma_{j,k} dL_k(u)\right)\right] \\
&\quad = \sum_{k=1}^n \sigma_{i,k}\sigma_{j,k} \int_0^\infty e^{-\lambda_j(t+u)} e^{-\lambda_i u} du \\
&\quad = \frac{e^{-\lambda_j t}}{\lambda_i + \lambda_j} \sum_{k=1}^n \sigma_{i,k}\sigma_{j,k} \\
&\quad = ((\Lambda_{var} \circ \Sigma_L \Sigma_L^\top) e^{-\Lambda t})_{i,j}
\end{aligned}$$

from which (C.10) follows. □

We now turn to proving Prop. 3.2. Initially, we give the next result which is a special case of (Benth & Rohde, 2019, Theorem 4.8), but again, we give a proof for the sake of completeness.

Proposition A.2. *Let L be a compound Poisson process with intensity $\alpha > 0$ and exponential jumps with parameter $\beta > 0$. Consider $t \in \mathbb{R}$, $\lambda, \mu > 0$ and $x_1, x_2 \in \mathbb{R}$ with $x_1 + x_2 < \beta$. Furthermore, assume $x_1 x_2 \geq 0$ and $x_1 \neq 0$, and let $f(t) =$*

$x_1 e^{-\lambda t} + x_2 e^{-\mu t}$. *Then*

$$\log \mathbb{E}\left[\exp\left\{\int_{-\infty}^{t} f(t-u)dL(u)\right\}\right] = \alpha \frac{f(0)}{f'(0)} \log\left(1 - \frac{f(0)}{\beta}\right) + \alpha \int_0^\infty \left(\frac{f(u)}{f'(u)}\right)' \log\left(1 - \frac{f(u)}{\beta}\right) du, \tag{C.28}$$

where

$$\left|\left(\frac{f(u)}{f'(u)}\right)'\right| \leq \frac{(\lambda-\mu)^2}{2\lambda\mu} \tag{C.29}$$

for all $u \geq 0$.

Proof. Initially, note that f/f' is bounded and

$$\begin{aligned}\left|\left(\frac{f(u)}{f'(u)}\right)'\right| &= \frac{|f'(u)^2 - f(u)f''(u)|}{f'(u)^2} \\ &= \frac{x_1 x_2 (\lambda-\mu)^2 e^{-(\lambda+\mu)u}}{x_1^2\lambda^2 e^{-2\lambda u} + x_2^2\mu^2 e^{-2\mu u} + 2x_1x_2\lambda\mu e^{-(\lambda+\mu)u}} \\ &\leq \frac{(\lambda-\mu)^2}{2\lambda\mu}.\end{aligned}$$

This gives the bound on $(f/f')'$. Additionally, we find that

$$\begin{aligned}\left|\left(\frac{f(u)}{f'(u)}\right)'\right| &= \frac{x_1 x_2 (\lambda-\mu)^2 e^{-(\lambda+\mu)u}}{x_1^2\lambda^2 e^{-2\lambda u} + x_2^2\mu^2 e^{-2\mu u} + 2x_1x_2\lambda\mu e^{-(\lambda+\mu)u}} \\ &= \frac{x_1 x_2 (\lambda-\mu)^2}{x_1^2\lambda^2 e^{-(\lambda-\mu)u} + x_2^2\mu^2 e^{-(\mu-\lambda)u} + 2x_1x_2\lambda\mu}.\end{aligned}$$

We conclude that $(f(u)/f'(u))' = O(e^{-|\lambda-\mu|u})$ as $u \to \infty$. Thus, all integrals below are convergent and the integration by parts is justified. Next let

$$\psi(u) = \log \mathbb{E}[\exp(uL(1))] = \alpha \frac{u}{\beta - u}$$

be the cumulant-generating function of $L(1)$ and let $\phi(u) = -\alpha \log(1 - u/\beta)$ be the cumulant-generating funciton of a gamma distribution with shape α and rate β (see for example Bertsekas & Pircalabu (2018)). Note that $\psi(u) =$

$u\phi'(u)$. Then, using integration by parts,

$$\begin{aligned}
&\log \mathbb{E}\left[\exp\left\{\int_{-\infty}^{t} f(t-u)dL(u)\right\}\right] \\
&\quad = \int_0^\infty \psi(f(u))du \\
&\quad = \int_0^\infty \frac{f(u)}{f'(u)}(\phi(f(u)))'du \\
&\quad = -\frac{f(0)}{f'(0)}\phi(f(0)) - \int_0^\infty \left(\frac{f(u)}{f'(u)}\right)'\phi(f(u))du. \qquad \square
\end{aligned}$$

Remark A.3. *Considering the proof of Prop. A.2 there are two approaches to calculate*

$$\log \mathbb{E}\left[\exp\left\{\int_{-\infty}^{t} f(t-u)dL(u)\right\}\right]. \tag{C.30}$$

Either by calculating

$$\int_0^\infty \psi(f(u))du \tag{C.31}$$

or

$$-\frac{f(0)}{f'(0)}\phi(f(0)) - \int_0^\infty \left(\frac{f(u)}{f'(u)}\right)'\phi(f(u))du. \tag{C.32}$$

Here, ψ and ϕ are the cumulant-generating function of $L(1)$ and of a gamma distribution with shape α and rate β as defined in the proof of Prop. A.2. By (C.29)*, the integral in* (C.32) *will be small whenever $(\lambda-\mu)^2/(2\lambda\mu)$ is small. In the application we consider we are concerned with the case where $\lambda = \hat{\lambda}_i$ and $\mu = \hat{\lambda}_j$ for some $i, j = 1, 2, 3, 4$, where $\hat{\lambda}_i$ and $\hat{\lambda}_j$ are given in Table C.2. We have*

$$\max_{i,j} \frac{(\hat{\lambda}_i - \hat{\lambda}_j)^2}{2\hat{\lambda}_i\hat{\lambda}_j} = 0.0506,$$

and therefore, indeed, that $(\lambda-\mu)^2/(2\lambda\mu)$ is small in the case relevant to us. The integral in (C.32) *has ϕ in the kernel whereas* (C.31) *has ψ, making a direct comparison more difficult. We do, however, have*

$$\phi(u) = \alpha u + O(u^2) \text{ and } \psi(u) = \alpha u + O(u^2) \text{ as } u \to 0$$

(by a Taylor approximation argument), indicating that ϕ and ψ are of comparable size, at least for small values. Furthermore, by numerical comparison, we have found them to be of similar size. We conclude that the kernel of (C.32) *is expected to be considerably smaller than the kernel of* (C.31)*. We therefore prefer to do the calculation in* (C.32) *instead of* (C.31) *since we can do a much more coarse approximation for a desired precision of a approximation of* (C.30)*.*

Proposition A.4. *Let L be a compound Poisson process with intensity $\alpha > 0$ and exponential jumps with parameter $\beta > 0$. Consider $s < t$, $\lambda > 0$ and $x < \beta$. Then*

$$\mathbb{E}\left[\exp\left\{x\int_s^t e^{-\lambda(t-u)}dL(u)\right\}\right] = \left(\frac{\beta - xe^{-\lambda(t-s)}}{\beta - x}\right)^{\alpha/\lambda}$$

and

$$\mathbb{E}\left[\exp\left\{x\int_{-\infty}^t e^{-\lambda(t-u)}dL(u)\right\}\right] = \left(\frac{\beta}{\beta - x}\right)^{\alpha/\lambda} \tag{C.33}$$

Proof. Let

$$\psi(t) = \log\mathbb{E}[\exp(tL(1))] = \alpha\frac{t}{\beta - t}$$

be the cumulant-generating function of L. Then

$$\begin{aligned}
&\log\mathbb{E}\left[\exp\left\{\int_s^t f(t-u)dL(u)\right\}\right] \\
&\quad= \int_0^{t-s}\psi(e^{-\lambda u})du \\
&\quad= \frac{\alpha}{\lambda}\left(\log(\beta - xe^{-\lambda(t-s)}) - \log(\beta - x)\right).
\end{aligned}$$

A similar calculation gives (C.33). □

Proof of Theorem 3.2. For notional convenience, let

$$\tilde{L}(t) = (L_1(t), \ldots, L_{n-1}(t), 0, L_n(t))^\top \in \mathbb{R}^{n+1}.$$

First consider (C.7) and assume $i \neq j$. We have

$$\begin{aligned}
X_i(t) = &\int_{-\infty}^t e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \int_s^t e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u) \\
&+ \int_{-\infty}^s e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)
\end{aligned}$$

and

$$X_j(s) = \int_{-\infty}^s e^{-\lambda_j(s-u)}\tilde{\sigma}_{j,j}d\tilde{L}_j(u) + \int_{-\infty}^s e^{-\lambda_j(s-u)}\tilde{\sigma}_{j,n+1}d\tilde{L}_{n+1}(u).$$

Next, note that $\operatorname{cov}(UV, W) = \operatorname{cov}(V, UW) = \mathbb{E}[U]\operatorname{cov}(V, W)$ for a random variable U independent of the random variables V and W. Applying this,

and the above, we conclude that

$$\begin{aligned}
&\operatorname{cov}(P_i(t), P_j(s)) \\
&= \operatorname{cov}\left(e^{-S_i(t)X_i(t)}, e^{-S_j(s)X_j(s)}\right) \\
&= \mathbb{E}\left[\exp\left\{-S_i(t)\int_{-\infty}^{t} e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,i}d\tilde{L}_i(u)\right\}\right] \\
&\times \mathbb{E}\left[\exp\left\{-S_i(t)\int_{s}^{t} e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right\}\right] \\
&\times \mathbb{E}\left[\exp\left\{-S_j(s)\int_{-\infty}^{s} e^{-\lambda_j(t-u)}\tilde{\sigma}_{j,j}d\tilde{L}_j(u)\right\}\right] \\
&\times \operatorname{cov}\left(\exp\left\{-S_i(t)\int_{-\infty}^{s} e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,n+1}d\tilde{L}_i(u)\right\},\right. \\
&\qquad \left.\exp\left\{-S_j(s)\int_{-\infty}^{s} e^{-\lambda_j(s-u)}\sigma_{j,n+1}d\tilde{L}_{n+1}(u)\right\}\right)
\end{aligned} \tag{C.34}$$

Expressions of the three expectations in Eq. (C.34) are given in Prop. A.4. Furthermore,

$$\begin{aligned}
&\operatorname{cov}\left(\exp\left\{-S_i(t)\int_{-\infty}^{s} e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,n+1}d\tilde{L}_i(u)\right\},\right. \\
&\qquad \left.\exp\left\{-S_j(s)\int_{-\infty}^{s} e^{-\lambda_j(s-u)}\tilde{\sigma}_{j,n+1}d\tilde{L}_{n+1}(u)\right\}\right) \\
&= \mathbb{E}\left[\exp\left\{-\int_{-\infty}^{s} f_{i,j}(s-u)d\tilde{L}_{n+1}(u)\right\}\right] \\
&- \mathbb{E}\left[\exp\left\{-S_i(t)\int_{-\infty}^{s} e^{-\lambda_i(t-u)}\tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right\}\right] \\
&\times \mathbb{E}\left[\exp\left\{-S_j(s)\int_{-\infty}^{s} e^{-\lambda_j(s-u)}\tilde{\sigma}_{j,n+1}d\tilde{L}_{n+1}(u)\right\}\right]
\end{aligned}$$

for which expressions are given in Prop. A.2 and Prop. A.4.

Next, consider (C.8). We write

$$\begin{aligned}
X_i(t) = &\int_{s}^{t} e^{-\lambda_i(t-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right) \\
&+ \int_{-\infty}^{s} e^{-\lambda_i(t-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right)
\end{aligned}$$

and

$$X_i(s) = \int_{-\infty}^{s} e^{-\lambda_i(s-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right).$$

Consequently,

$$\begin{aligned}
&\operatorname{cov}(P_i(t), P_i(s)) \\
&\quad = \operatorname{cov}\left(e^{-S_i(t)X_i(t)}, e^{-S_i(s)X_i(s)}\right) \\
&\quad = \mathbb{E}\left[\exp\left\{-S_i(t)\int_s^t e^{-\lambda_i(t-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right)\right\}\right] \\
&\quad \times \operatorname{cov}\left(\exp\left\{-S_i(t)\int_{-\infty}^s e^{-\lambda_i(t-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right)\right\},\right. \\
&\qquad \left.\exp\left\{-S_i(s)\int_{-\infty}^s e^{-\lambda_i(s-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_j(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right)\right\}\right)
\end{aligned}$$

Again, expressions for the expectation in (C.34) can be found using Prop. A.4. Finally,

$$\begin{aligned}
&\operatorname{cov}\left(\exp\left\{-S_i(t)\int_{-\infty}^s e^{-\lambda_i(t-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right)\right\},\right. \\
&\qquad \left.\exp\left\{-S_i(s)\int_{-\infty}^s e^{-\lambda_i(s-u)}\left(\tilde{\sigma}_{i,i}d\tilde{L}_i(u) + \tilde{\sigma}_{i,n+1}d\tilde{L}_{n+1}(u)\right)\right\}\right) \\
&\quad = \mathbb{E}\left[\exp\left\{-\int_{-\infty}^s \tilde{\sigma}_{i,i}\left(S_i(t)e^{-\lambda_i(t-s)} + S_i(s)\right)e^{-\lambda_i(s-u)}d\tilde{L}_i(u)\right\}\right] \\
&\quad \times \mathbb{E}\left[\exp\left\{-\int_{-\infty}^s \tilde{\sigma}_{i,n+1}\left(S_i(t)e^{-\lambda_i(t-s)} + S_i(s)\right)e^{-\lambda_i(s-u)}d\tilde{L}_{n+1}(u)\right\}\right] \\
&\quad - \mathbb{E}\left[\exp\left\{-\int_{-\infty}^s \tilde{\sigma}_{i,i}S_i(t)e^{-\lambda_i(t-u)}d\tilde{L}_i(u)\right\}\right] \\
&\quad \times \mathbb{E}\left[\exp\left\{-\int_{-\infty}^s \tilde{\sigma}_{i,n+1}S_i(t)e^{-\lambda_i(t-u)}d\tilde{L}_{n+1}(u)\right\}\right] \\
&\quad \times \mathbb{E}\left[\exp\left\{-\int_{-\infty}^s \tilde{\sigma}_{i,i}S_i(s)e^{-\lambda_i(s-u)}d\tilde{L}_i(u)\right\}\right] \\
&\quad \times \mathbb{E}\left[\exp\left\{-\int_{-\infty}^s \tilde{\sigma}_{i,n+1}S_i(s)e^{-\lambda_i(s-u)}d\tilde{L}_{n+1}(u)\right\}\right]
\end{aligned}$$

where expressions are given in Prop. A.4. □

References

Asmussen, S., Jensen, J. L., Rojas-Nandayapa, L. (2016). On the Laplace transform of the lognormal distribution. *Methodology and Computing in Applied Probability*, **18**, pp. 441–458.

Ballotta, L., Bonfiglioli, E. (2016). Multivariate asset models using Lévy processes and applications. *The European Journal of Finance*, **22(13)**, pp. 1320–1350.

Barndorff-Nielsen, O. E., Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **63(2)**, pp. 167–241.

Benth, F. E., Benth, J. Š. (2007). The volatility of temperature and pricing of weather derivatives. *Quantitative Finance*, **7(5)**, pp. 553-561.

Benth, F. E., Kallsen, J., Meyer-Brandis, T. (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, **14(2)**, pp. 153–169.

Benth, F. E., Pircalabu, A. (2018). A non-Gaussian Ornstein–Uhlenbeck model for pricing wind power futures. *Applied Mathematical Finance*, **25**, pp. 36–65.

Benth, F. E., Rohde, V. (2019). On non-negative modeling with CARMA processes. *Journal of Mathematical Analysis and Applications*, **476(1)**, pp. 196–214.

Christensen, T. S., Pircalabu, A. (2018). On the spatial hedging effectiveness of German wind power futures for wind power generators. *Journal of Energy Markets*, **11**, pp. 71–96.

Christensen, T. S., Pircalabu, A., Høg, E. (2019). A seasonal copula mixture for hedging the clean spark spread with wind power futures *Energy Economics*, **78**, pp. 64–80.

Gersema, G. and Wozabal, D. (2017). An equilibrium pricing model for wind power futures. *Energy Economics*, **65**, pp. 64–74.

Halliwell, L. J. (2015). The lognormal random multivariate. *Casualty Actuarial Society E-Forum*.

Leoni, P., Schoutens, W. (2008). Multivariate Smiling. *Wilmott Magazines*.

Meucci, A. (2009). Review of statistical arbitrage, cointegration, and multivariate Ornstein-Uhlenbeck. *Available at SSRN: https://ssrn.com/abstract=1404905*.

Rajput, B. S., Rosinski, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, **82(3)**, pp. 451–487.

Sato, K. I. (1999). Lévy processes and infinitely divisible distributions. *Cambridge University Press.*.

Semeraro, P. (2008). A multivariate variance gamma model for financial application. *International Journal of Theoretical and Applied Finance*, **11(1)**, pp. 1–18.

Paper D

Modeling the Joint Behaviour of Electricity Prices in Interconnected Markets

List of authors: Troels Sønderby Christensen[1,2], Fred Espen Benth[3]

[1]Department of Mathematical Sciences, Aalborg University
[2]Quantitative Analytics, Centrica Energy Trading
[3]Department of Mathematics, University of Oslo

The paper has been submitted to Quantitative Finance.

The layout has been revised.

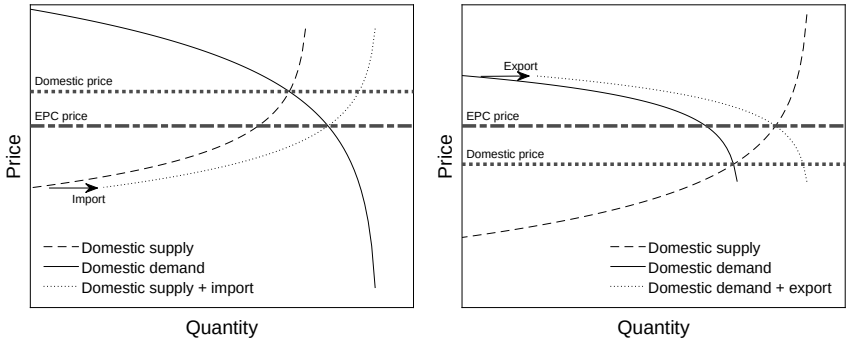
Abstract

The liberalization of energy markets worldwide during recent decades has introduced severe implications on the price formation in these markets. Especially within the European day-ahead electricity markets, increased physical connections between different market areas and a joint effort on optimizing the aggregate social welfare have led to highly connected markets. Consequently, observing the exact same hourly day-ahead prices for two or more interconnected electricity markets in Europe happens frequently. This affects the modeling of such prices and in turn the valuation of derivatives written on prices from such market areas. In this paper, we propose a joint model for day-ahead electricity prices in interconnected markets composed of a combination of transformed Ornstein-Uhlenbeck processes. We discuss the properties of the model and propose an estimation procedure based on filtering techniques. Furthermore, the properties of the model reveal that analytical prices are attainable for e.g. forwards and spread options.

1 Introduction

With the introduction of liberalized energy markets worldwide, increased interconnectedness between different market areas has emerged. The so-called market coupling mechanism has increased the degree of connectivity even more in the European electricity markets. The market coupling works by collecting bids and offers for trading electricity the following day from each market area. Then, based on a social welfare optimization, a single algorithm calculates optimal flows in the system given interconnector limitations for each hour of the following day and the corresponding electricity prices; hence, the name day-ahead prices (Epex Spot (2019)). Fig. D.1 illustrates this procedure in a two-market setup. The prices labelled "Domestic price" are the prices set if the areas were not connected and no cross-border flows of electricity were present. The difference between the domestic prices are rather large in this case. Though, by the market coupling mechanism, the price spread between these two areas are minimized. In the case of Fig. D.1, we have enough interconnector capacity between the areas to eliminate the price spread completely. We define such a zero price spread as exact price convergence (EPC). The result is that cheap electricity is transferred from the low price area to the high price area, or loosely speaking, electricity flows from the area with a surplus of electricity to the area with a need for electricity.

As exemplified by Fig. D.1, the market coupling has severe implications on the resulting prices in interconnected markets. Most notable is the presence of EPC that naturally has an impact on derivative prices, and in particular on *transmission rights*. A transmission right is basically a spread option between two connected market areas with the underlying being the spread



(a) High price area (b) Low price area

Fig. D.1: Example of market coupling in a two-market setup where exact price convergence is reached; i.e., the price spread between the two market areas is forced to be *exactly* zero by the market coupling mechanism. As a result, the domestic price in the high price area is reduced and the domestic price in the low price area is increased.

between the market areas under consideration[1]. When EPC occurs, the underlying spread is zero, and the transmission right becomes worthless. To conduct sensible pricing of transmission rights (and related derivatives in interconnected markets), a model capable of incorporating zero price spreads is therefore motivated.

Day-ahead electricity price modeling in general has been examined in numerous studies. To name a few, Lucia & Schwartz (2002), Geman & Roncoroni (2006), and Meyer-Brandis & Tankov (2008) all focus on modeling the temporal dynamics of a single market. Lucia & Schwartz (2002) exemplify their proposed model using daily average prices from the Nord Pool market, Geman & Roncoroni (2006) use daily prices from the US, while Meyer-Brandis & Tankov (2008) use EEX prices. Benth & Šaltytė Benth (2004) and Benth *et al.* (2007) also focus on the temporal dynamics of a single market, and additionally discuss derivative pricing. Both of them present the model in the context of Nord Pool data, while the former also uses Brent oil data. Regarding the spatial interdependencies between different market areas, we mention Douglas & Popova (2011) and Abate & Haldrup (2017). Douglas & Popova (2011) estimate a spatial error model on data from Pennsylvania-New Jersey-Maryland Interconnection, consisting of twelve market areas at the time of their investigation. Abate & Haldrup (2017) investigate daily averages of day-ahead prices from the markets making up the Nord Pool area, and find that the inclusion of spatial effects is as important as including tem-

[1]A transmission right can either be physically or financially settled, and will collapse if the physical transmission right is sold under the "use it or sell it" condition. We refer the interested reader to Mahringer *et al.* (2015) for more information.

poral effects.

The literature concerning the joint modeling of day-ahead electricity prices for two or more interconnected markets where market coupling is taken into account is quite scarce. Mahringer *et al.* (2015) use a structural approach, where they model the supply and demand in a two-market setup that incorporates market coupling and allows for closed-form solutions for transmission rights. Pircalabu & Benth (2017) propose a reduced-form model in a two-market setup defined in discrete time that includes the market coupling mechanism using a regime-switching model, but no closed-form solutions for derivative prices are given.

As Pircalabu & Benth (2017) we propose a reduced-form model for day-ahead prices in a two-market setup that takes market coupling into account. Compared to Pircalabu & Benth (2017), we differentiate us by formulating the model in continuous time. Instead of an ARMA-GARCH copula formulation combined with a discrete Markov chain, we assume that day-ahead prices can be described by a suitable combination and transformation of Gaussian Ornstein-Uhlenbeck processes. On one hand, this has the disadvantage that complicated dependence structures that are easily handled by copulas can be rather hard to incorporate in the present continuous-time formulation. On the other hand, the continuous-time formulation makes us able to derive closed-form solutions of forward and transmission right prices. Our contribution to the literature additionally includes an estimation procedure for the proposed model. It is based on two independent parts, namely a Kalman filtering part and a particle filtering part.

Although we in this paper have a Eurpoean perspective, we believe that the results are interesting in other parts of the world as well where market area dependencies are present. As an example, Park *et al.* (2006) examine connected electricity markets in the US, and find that day-ahead prices in such connected market areas inherit spatial interdependencies. Also interesting is the work in Ignatieva & Trück (2016), studying the dependence in day-ahead prices between different market areas in the Australian electricity market. Ignatieva & Trück (2016) find a positive dependence between all market areas, but the strongest dependence is observed between adjacent market areas.

The remaining part of this paper is organized as follows: Sec. 2 introduces data and also serves as a motivation for the model choices made in later sections. In Sec. 3 we introduce the proposed model and discuss its properties. Sec. 4 concerns estimation of the model, while we in Sec. 5 exemplify the model empirically through a case study of German and French electricity prices and discuss derivative prices implied by the model. Sec. 6 concludes.

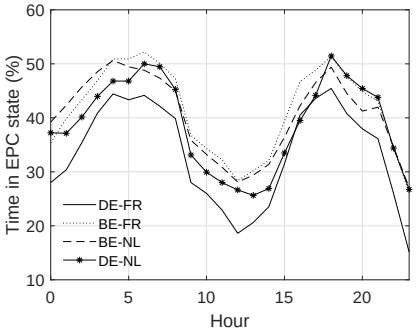


Fig. D.2: Time spent in the EPC state in percentage for each hour from 1 July 2015 to 30 June 2018 for the following spreads: DE-FR, BE-FR, BE-NL, and DE-NL.

2 Data Presentation and Model Motivation

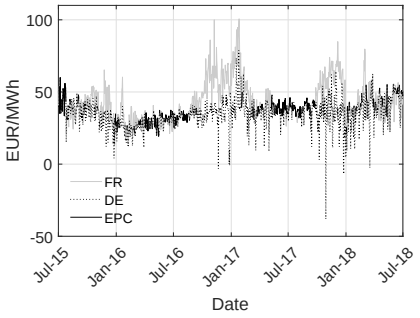
This section presents data from the hourly power markets under consideration, and motivates the model choices made in later sections. The observations in this paper cover the three-year period from 1 July 2015 to 30 June 2018 and the market areas making up the Central Western Europe (CWE) area; Germany, France, the Netherlands, and Belgium. The market coupling algorithm was actually already introduced in the CWE area in November 2010; though, the underlying optimization algorithm was changed from the so-called available transfer capacity method to the flow-based market coupling method on 20 May 2015 (see Elia (2019)). In order to avoid results affected by the old methodology, we have thus chosen to fix our start date to 1 July 2015.

To give a sense of the frequency of how often EPC happens, Table D.1 reports the observed percentage of time EPC occurred for each hour during the mentioned three years. The EPC percentage varies quite a bit depending on the chosen price spread and chosen hour. Fig. D.2 illustrates the data showed in Table D.1, and shows a clear pattern throughout the day for all spreads, with peaks around hour 5 and 18. The lowest values are reached at hour 12 and 23. Though, the rather high percentages in general suggests that EPC is important to consider in a modeling context in all cases.

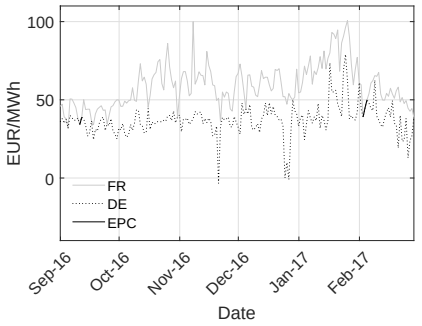
Fig. D.3(a) illustrates the evolution of the day-ahead electricity price in Germany and France for the hour from 21:00-22:00. In the rest of the paper, we denote the hour spanning the period 21:00-22:00 by hour 21 and likewise for the other hours. First of all, negative prices are observed, indicating oversupply of electricity in the grid. Traditional financial models that only allow non-negative prices are therefore not able to take this feature into account. Even though it is only German prices that are negative for this particular time series, negative values are allowed for the other CWE market areas as

Hour	DE-FR	BE-FR	BE-NL	DE-NL
0	28	35	39	37
1	30	40	42	37
2	35	43	46	40
3	41	47	49	44
4	44	51	51	47
5	43	51	49	47
6	44	52	49	50
7	42	50	47	49
8	40	47	45	45
9	28	37	36	33
10	26	34	33	30
11	23	32	31	28
12	19	28	28	27
13	21	30	29	26
14	24	32	31	27
15	31	39	36	33
16	41	47	42	40
17	44	49	47	44
18	45	51	49	51
19	41	48	45	48
20	38	45	41	45
21	36	43	42	4
22	26	35	35	34
23	15	26	27	27

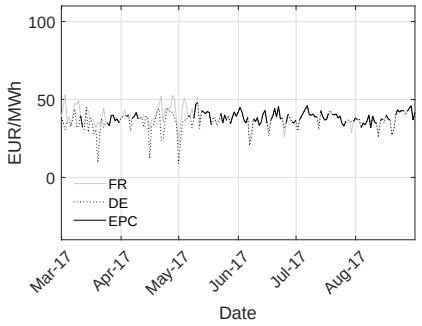
Table D.1: Observed EPC percentage for each hour for the four different cross-border connections in Central-West Europe from 1 July 2015 to 30 June 2018.



(a) Day-ahead prices for the entire period



(b) Sub sample of autumn/winter day-ahead prices

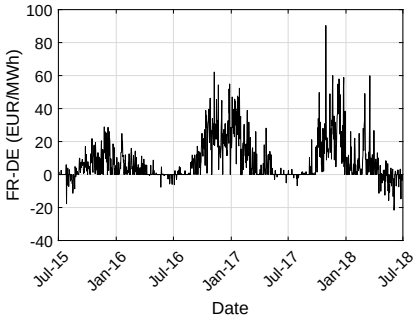


(c) Sub sample of spring/summer day-ahead prices

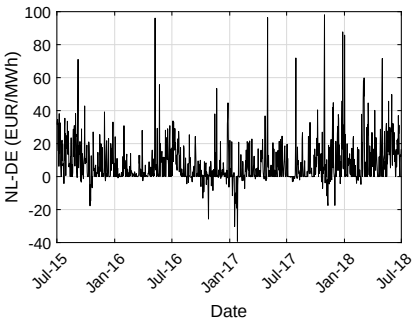
Fig. D.3: (a) Day-ahead pricse for France and Germany for hour 21 from 1 July 2015 to 30 June 2018, (b) and (c) are subsamples of the data presented in Fig. D.3(a) to illustrate the differences between these periods.

well. Secondly, it is apparent that the mean and volatility of each price series seems to be highly seasonal. In particular, spring and summer periods result in a more calm behaviour, and autumn and winter periods in a more volatile behaviour, see Figs. D.3(b)–D.3(c) that exemplify the seasonal differences.

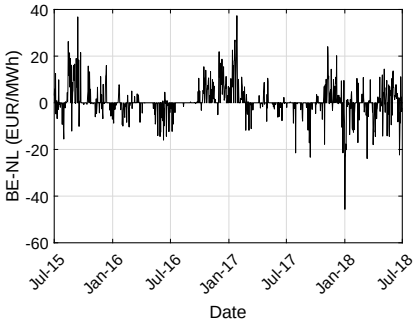
Proceeding with the seasonal behaviour, Fig. D.4 illustrates the day-ahead price differences between three chosen price pairs. We see that EPC happens more often during summer and spring than during autumn and winter in the FR-DE case in Fig. D.4(a). A possible explanation for this is the fact that the cheap nuclear power production from France during summer levels out the differences between the French and German prices. Opposite, during winter the demand for electricity rises in France and in turn the French power prices rise, while the German prices do not rise similarly partly due to the high wind power production and the lower dependency on using electricity for heating, all in all making the gap between German and French prices widen during autumn/winter. With a physical connection between Germany



(a) FR-DE hour 21 day-ahead price spread



(b) NL-DE hour 12 day-ahead price spread



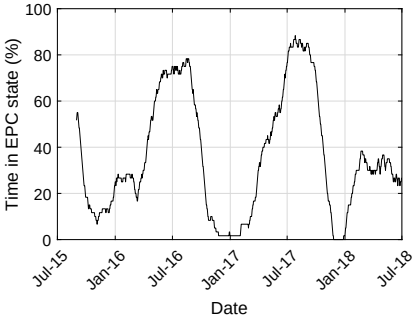
(c) BE-NL hour 8 day-ahead price spread

Fig. D.4: Day-ahead price differences for different hours and market areas.

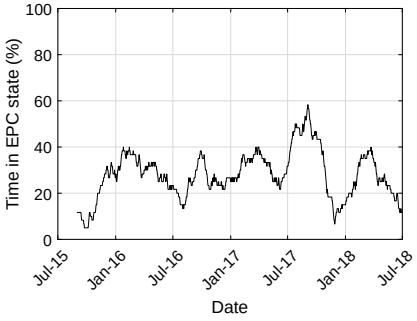
and France being only of a certain size, and since the amount of needed transferred electricity between Germany and France to level out prices during autumn/winter is greater due to the larger price gap, congestion is observed more often during autumn/winter than spring/summer. A similar behaviour is observed in Fig. D.4(c), while the rate at which EPC occurs seems more evenly spread out over the year in Fig. D.4(b).

To investigate EPC in more detail, we plot in Fig. D.5 the number of days being in the EPC state for the spreads considered in Fig. D.4. The plot is constructed by dividing the observations in overlapping sub-samples of 60 days, and then calculating the number of EPC days in each sub sample. Fig. D.5(a) shows a rather large difference between the time spent in EPC during spring/summer and autumn/winter, with a peak above 80% and the lowest value being 0%. The same seasonal pattern is observed in Fig. D.5(c), though with a lower amplitude. Opposite, the level is much more constant in the case of Fig. D.5(b).

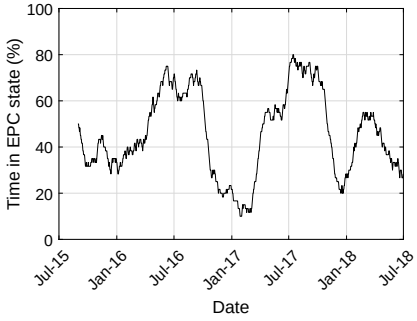
As is apparent from the above discussion, special features distinguish hourly day-ahead prices from traditional financial markets. Besides the ob-



(a) FR-DE hour 21 day-ahead price



(b) NL-DE hour 12 day-ahead price



(c) BE-NL hour 8 day-ahead price

Fig. D.5: Time spent in the EPC state in percentage for overlapping sub-samples of 60 days.

vious allowed negativity in the CWE area, seasonality in the mean level and volatility are known features of electricity prices (see e.g. Haldrup *et al.* (2010), Pircalabu & Benth (2017), and Borovkova & Schmeck (2017)), but as discussed above time-varying seasonality in probability of being in the EPC state is also important to take into account.

3 Model Description

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We propose the following model for the day-ahead prices

$$S(t) = I(t)p(t), \tag{D.1}$$

where $S(t) \in \mathbb{R}^N$ is the vector of day-ahead prices at time t for N market areas. The motivation for our model is to capture the price formation mechanism in a reduced-form manner. The exact convergence of day-ahead prices is represented by $I(t) \in \mathbb{R}^{N\times N}$, which is piecewise constant resembling the exact price convergence of prices from different market areas. An illustrative example of the process $I(t)$ is in the case $N = 2$, where one can choose the states of $I(t)$ as

$$I_{(1)} = \begin{bmatrix} w & 1-w \\ w & 1-w \end{bmatrix} \text{ and } I_{(2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{D.2}$$

Here w is a constant between zero and one that measures the pulling effect of each component of $S(t)$. In this example, state two—emphasized with subscript (2)—corresponds to no exact price convergence and $S(t) = p(t)$, whereas state one corresponds to exact price convergence with the common price being $wp_1(t) + (1-w)p_2(t)$. The states of $I(t)$ is driven by an underlying process $z(t)$, which we will discuss shortly.

The process $p(t) \in \mathbb{R}^N$ represents the partly latent non-convergence price set in each individual market area, and can thus be interpreted as the "domestic" prices. It is only partly latent, since we do in fact observe it in non EPC states. In the above example with $N = 2$, we observe $p(t)$ whenever we are in state two. Based on the rather deterministic demand for electricity (which in turn is due to highly seasonal temperature patterns influencing the demand to a large extent) we assume that deviations from a deterministic seasonal function are temporary. We therefore assume that $p(t)$ can be divided in a deterministic part, $\Lambda(t) \in \mathbb{R}^N$ and stochastic part $q(t) \in \mathbb{R}^N$, by

$$p(t) = \Lambda(t) + q(t). \tag{D.3}$$

We assume that the stochastic part is governed by a multivariate Ornstein-Uhlenbeck process

$$dq(t) = -Kq(t)dt + \Sigma dB_p(t), \tag{D.4}$$

where $K \in \mathbb{R}^{N\times N}$ is a constant matrix with the real part of its eigenvalues being positive, $\Sigma \in \mathbb{R}^{N\times N}$, and $B_p(t)$ is an N-dimensional Brownian motion. By this specification we allow for negative prices. While this is clearly unsatisfactory in traditional financial markets, negative prices in electricity markets do occur as discussed in Sec. 2.

We remark that a similar modelling idea has been proposed in a recent paper by Hinderks *et al.* (2018) in a structural Heath-Jarrow-Morton approach to electricity futures pricing. There, fixed-delivery futures price dynamics are modelled as a product of a noise process and a conditional expectation of a structural component. As for our model, the authors demonstrate the analytical tractability of such an approach in derivatives pricing.

The above specification holds for all N; however, we will limit ourselves to the case where $N = 2$ in the rest of the paper for two reasons. First of all, an important feature of the model is the ability to price transmission rights. By construction, these derivatives are essentially spread options with the underlying being the spread between only *two* adjacent market areas. We thus argue that the case with $N > 2$ is unnecessary regarding this application, but acknowledge that it could be relevant in other contexts. Secondly, modeling issues related to the underlying process governing $I(t)$ arises when $N > 2$. We briefly comment on this in the next sections, where we elaborate on the components of the model in more detail.

3.1 Properties of $p(t)$

Let $\text{vec}(\cdot)$ denote the vectorization operator that stacks the columns of a matrix A on top of each other; i.e., if A is an $m \times n$ matrix, then the vectorization operator applied on A results in a column vector of size $mn \times 1$. Let further $\text{vec}^{-1}(\cdot)$ be the operator satisfying $\text{vec}^{-1}(\text{vec}(A))=A$ for a square matrix A. Notice that for non-square matrices, $\text{vec}^{-1}(\cdot)$ is not uniquely defined; however, since we only deal with square matrices, we omit a further discussion of this. Define the Kronecker sum between two square matrices A and B by $A \oplus B = A \otimes \mathbf{I}_B + \mathbf{I}_A \otimes B$, where $\otimes$ is the Kronecker product, and $\mathbf{I}_B$ is the identity matrix with the same size as B, and $\mathbf{I}_A$ is defined in the same way. We denote the filtration generated by the Brownian motion $B_p(t)$ by $\{\mathcal{F}_t^{B_p}\}_{t\geq 0}$. Prop. 3 states the distribution of $p(t)$ conditional on $\mathcal{F}_s^{B_p}$.

Proposition 3
Let $p(t)$ be defined as in Eq. (D.3), and let $s < t$. The distribution of $p(t)$ is normal conditional on $\mathcal{F}_s^{B_p}$ with mean

$$\mathbb{E}[p(t)|\mathcal{F}_s^{B_p}] = \Lambda(t) + e^{-(t-s)K}q(s) \tag{D.5}$$

and variance

$$\mathbb{V}\mathrm{ar}[p(t)|\mathcal{F}_s^{B_p}] = \mathrm{vec}^{-1}\left((K \oplus K)^{-1}(\mathbf{I} - e^{-(t-s)K\oplus K})\mathrm{vec}(\Sigma\Sigma^{\mathrm{T}})\right), \tag{D.6}$$

where $\mathbf{I}$ is the identity matrix of same size as $K \oplus K$.

Proof. The solution of Eq. (D.4) is seen to be

$$q(t) = e^{-(t-s)K}q(s) + \int_s^t e^{-(t-u)K}\Sigma dB_p(u) \tag{D.7}$$

by the multidimensional Ito formula. The normality and expected value follows. Using the Ito isometry, we get

$$\mathbb{V}\mathrm{ar}[p(t)|\mathcal{F}_s^{B_p}] = \int_s^t e^{-(t-u)K}\Sigma\Sigma^{\mathrm{T}}e^{-(t-u)K^{\mathrm{T}}}du. \tag{D.8}$$

For matrices A, B, and C of appropriate sizes, two useful identities are $\mathrm{vec}(ABC) = (C^{\mathrm{T}} \otimes A)\mathrm{vec}(B)$ (see e.g. Macedo & Oliveira (2013) and the references herein for more information on this result) and $e^{A\oplus B} = e^A \otimes e^B$. Eq. (D.8) can then be simplified by writing the vectorized version as

$$\begin{aligned}\mathrm{vec}\left(\mathbb{V}\mathrm{ar}[p(t)|\mathcal{F}_s^{B_p}]\right) &= \left(\int_s^t e^{-(t-u)(K\oplus K)}du\right)\mathrm{vec}(\Sigma\Sigma^{\mathrm{T}}) \\ &= (K \oplus K)^{-1}(\mathbf{I} - e^{-(t-s)K\oplus K})\mathrm{vec}(\Sigma\Sigma^{\mathrm{T}}).\end{aligned}$$

□

Since we assume that the eigenvalues of K all have positive real parts, all eigenvalues of $K \oplus K$ have positive real parts, see e.g. Laub (2005). Hence, the existence of a limiting distribution of $q(t)$ follows from Prop. 6.2 of Ichihara & Kunita (1974). This limiting distribution is normal with mean zero and variance

$$\lim_{t\to\infty}\mathbb{V}\mathrm{ar}[q(t)|\mathcal{F}_s^{B_p}] = \mathrm{vec}^{-1}\left((K \oplus K)^{-1}\mathrm{vec}(\Sigma\Sigma^{\mathrm{T}})\right). \tag{D.9}$$

As time approaches infinity, the expected value of $p(t)$ will therefore coincide with $\Lambda(t)$, while the variance of $p(t)$ coincides with the variance of $q(t)$ given in Eq. (D.9).

3.2 Properties of $z(t)$

Limiting our investigation to the case $N = 2$, we assume that the state of $I(t)$ is determined by a latent univariate stochastic process $z(t)$. Specifically, we assume that the EPC state occurs whenever $z(t) > 0$, so we in this two-dimensional case have

$$I(t) = \begin{cases} I_{(1)}, & \text{if } z(t) > 0 \\ I_{(2)}, & \text{otherwise} \end{cases},$$

where $I_{(1)}$ and $I_{(2)}$ are given in Eq. (D.2). $z(t)$ is defined by

$$z(t) = \lambda(t) + x(t), \tag{D.10}$$

$$\lambda(t) = a_\lambda + b_\lambda \cos\left(\frac{2\pi t}{M} + c_\lambda\right),$$

$$dx(t) = -kx(t)dt + \sigma(t)dB_z(t), \tag{D.11}$$

$$\sigma(t) = a_\sigma + b_\sigma \cos\left(\frac{2\pi t}{M} + c_\sigma\right).$$

Here a_λ, b_λ, c_λ, k, a_σ, b_σ, $c_\sigma, M \in \mathbb{R}$ are constants, and $B_z(t)$ is a univariate Brownian motion. $\lambda(t)$ is a seasonal function measuring the asymptotic mean level of $z(t)$, while $x(t)$ measures the uncertainty of $z(t)$. To allow for time-varying uncertainty in $\mathbb{P}(z(t) > 0)$, which is pronounced in the empirical investigations in Sec. 2 and especially Fig. D.5, we introduce seasonality in the volatility parameter $\sigma(t)$. We denote the filtration generated by the Brownian motion $B_z(t)$ by $\{\mathcal{F}_t^{B_z}\}_{t\geq 0}$, and assume independence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$. Define $\mathcal{F}_t := \mathcal{F}_t^{B_z} \vee \mathcal{F}_t^{B_p}$. The following proposition is useful in an estimation and simulation setting.

Proposition 4
Let $z(t)$ be given as in Eq. (D.10), and let $s < t$. Then the distribution of $z(t)$ is normal conditional on $\mathcal{F}_s$ with mean

$$\mathbb{E}[z(t)|\mathcal{F}_s] = \lambda(t) + x(s)e^{-k(t-s)}, \tag{D.12}$$

and variance

$$\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s] = \frac{a_\sigma^2}{2k}\left(1 - e^{-2k(t-s)}\right) + A(s,t) + B(s,t), \tag{D.13}$$

where

$$\begin{aligned} A(s,t) =& \frac{b_\sigma^2}{4(M^2k^3 + 4\pi^2 k)}\Bigg[4\pi^2 + M^2k^2 + M^2k^2\cos\left(\frac{4\pi t}{M} + 2c_\sigma\right) \\ &+ 2Mk\pi\sin\left(\frac{4\pi t}{M} + 2c_\sigma\right) - e^{-2k(t-s)}\Bigg(4\pi^2 + M^2k^2 \\ &+ M^2k^2\cos\left(\frac{4\pi s}{M} + 2c_\sigma\right) + 2Mk\pi\sin\left(\frac{4\pi s}{M} + 2c_\sigma\right)\Bigg)\Bigg], \end{aligned} \tag{D.14}$$

$$\begin{aligned} B(s,t) =& \frac{a_\sigma b_\sigma M}{M^2k^2 + \pi^2}\Bigg[\pi\sin\left(\frac{2\pi t}{M} + c_\sigma\right) + Mk\cos\left(\frac{2\pi t}{M} + c_\sigma\right) \\ &- e^{-2k(t-s)}\left(\pi\sin\left(\frac{2\pi s}{M} + c_\sigma\right) + Mk\cos\left(\frac{2\pi s}{M} + c_\sigma\right)\right)\Bigg]. \end{aligned} \tag{D.15}$$

Proof. Using the Ito formula, the solution to Eq. (D.11) for $s < t$ is

$$x(t) = e^{-k(t-s)}x(s) + \int_s^t e^{-k(t-u)}\sigma(u)dB_z(u).$$

The normality follows from the deterministic integrand and the Brownian motion. The expected value is trivially calculated, and the variance is obtained by using the Ito isometry and straightforward calculations. □

An Alternative Approach

One approach that might be obvious to use to model the underlying process governing the state of $I(t)$ is the continuous-time Markov chain, having showed to be of great usefulness in various applications (see e.g. Hubbard *et al.* (2008) in the context of disease progression modelling, and Inamura (2006) in the context of credit rating modelling). As already noted, the probability of EPC can be time dependent with a seasonal pattern, see Fig. D.5, meaning that a time-homogeneous Markov chain would fail to give a realistic representation of the underlying of $I(t)$. Time-inhomogeneous Markov chains would take care of the just-mentioned shortcoming, but, to the best of our knowledge, limited theoretical results exist to accommodate it. Hubbard *et al.* (2008) use a common time transformation to make transition rates time dependent, which is not applicable in our setting, since we would need a non-common time transformation to make certain states more likely than others in certain periods and vice versa. Time-inhomogeneous Markov chains being able to accommodate the above are therefore left as future research, which would possibly also pave the way for the case with $N > 2$.

3.3 Properties of $S(t)$

Having established results regarding $p(t)$ and $z(t)$ in Sections 3.1 and 3.2, respectively, we now turn to $S(t)$.

Proposition 5
Let $p(t)$ and $z(t)$ be defined as in Sections 3.1 and 3.2, respectively. The distribution of $S(t)$ is then mixed normal conditional on $\mathcal{F}_s$ for $s < t$ with density

$$f_{s,t}(x) = \sum_{j=1}^{2} \mathbb{P}_{(j)}(s,t)\frac{1}{(2\pi)^{N_j/2}|\Sigma_{(j)}(s,t)|}e^{-\frac{1}{2}\left(x-\mu_{(j)}(s,t)\right)^{\mathrm{T}}\Sigma_{(j)}(s,t)^{-1}\left(x-\mu_{(j)}(s,t)\right)}$$

with $N_j \leq N$ for $j = \{1,2\}$ being the rank of $I_{(j)}$, $\Phi(\cdot)$ being the cumulative distribution function for the standard normal distribution, and where

$$\begin{aligned}\mu_{(j)}(s,t) =& I_{(j)}\mathbb{E}[p(t)|\mathcal{F}_s]\\ \Sigma_{(j)}(s,t) =& I_{(j)}\mathbb{V}\mathrm{ar}[p(t)|\mathcal{F}_s]I_{(j)}^{\mathrm{T}},\end{aligned}$$

and

$$\mathbb{P}_{(1)}(s,t) = 1 - \Phi\left(\frac{-\mathbb{E}[z(t)|\mathcal{F}_s]}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right), \quad \mathbb{P}_{(2)}(s,t) = 1 - \mathbb{P}_{(1)}(s,t). \quad \text{(D.16)}$$

Proof. The characteristic function of $S(t)$ is

$$\begin{aligned}
\mathbb{E}\left[e^{iy^{\mathrm{T}}S(t)}|\mathcal{F}_s\right] =& \mathbb{E}\left[e^{iy^{\mathrm{T}}I(t)(\Lambda(t)+q(t))}|\mathcal{F}_s\right] \\
=& \mathbb{E}\Bigg[\mathbb{E}\Bigg[\exp\Bigg(iy^{\mathrm{T}}I(t)\Big(\Lambda(t) + e^{-(t-s)K}q(s) \\
&+ \int_s^t e^{-(t-u)K}\Sigma dB_p(u)\Big)\Bigg)|\mathcal{F}_t^{B_z} \vee \mathcal{F}_s\Bigg]|\mathcal{F}_s\Bigg] \\
=& \mathbb{E}\Bigg[\exp\left(iy^{\mathrm{T}}I(t)\left(\Lambda(t) + e^{-(t-s)K}q(s)\right)\right. \\
&\left. - \frac{1}{2}y^{\mathrm{T}}I(t)\int_s^t e^{-(t-u)K}\Sigma\Sigma^{\mathrm{T}}e^{-(t-u)K^{\mathrm{T}}}du I(t)^{\mathrm{T}}y\right)|\mathcal{F}_s\Bigg] \\
=& \mathbb{P}(z(t) > 0|\mathcal{F}_s)\exp\left(iy^{\mathrm{T}}I_{(1)}\left(\Lambda(t) + e^{-(t-s)K}q(s)\right)\right. \\
&\left. - \frac{1}{2}y^{\mathrm{T}}I_{(1)}\int_s^t e^{-(t-u)K}\Sigma\Sigma^{\mathrm{T}}e^{-(t-u)K^{\mathrm{T}}}du I_{(1)}^{\mathrm{T}}y\right) \\
&+ (1 - \mathbb{P}(z(t) > 0|\mathcal{F}_s))\exp\left(iy^{\mathrm{T}}I_{(2)}\left(\Lambda(t) + e^{-(t-s)K}q(s)\right)\right. \\
&\left. - \frac{1}{2}y^{\mathrm{T}}I_{(2)}\int_s^t e^{-(t-u)K}\Sigma\Sigma^{\mathrm{T}}e^{-(t-u)K^{\mathrm{T}}}du I_{(2)}^{\mathrm{T}}y\right).
\end{aligned}$$

The second equality is a consequence of the law of iterated expectations. As a result of Prop. 4, we further have that

$$\mathbb{P}(z(t) > 0|\mathcal{F}_s) = 1 - \Phi\left(\frac{-\mathbb{E}[z(t)|\mathcal{F}_s]}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right), \quad \text{(D.17)}$$

concluding the proof. □

The construction of the model for $S(t)$ implies that the introduced seasonal time variation in $p(t)$ and $z(t)$ translates to seasonal time variation in $S(t)$. To briefly elaborate on this, observe that the expected value of $S(t)$ is

$$\begin{aligned}
\mathbb{E}\left[S(t)|\mathcal{F}_s\right] &= \mathbb{E}\left[I(t)\mathbb{E}\left[p(t)|\mathcal{F}_t^{B_z} \vee \mathcal{F}_s\right]|\mathcal{F}_s\right] \\
&= \left[\mathbb{P}_{(1)}(s,t)I_{(1)} + (1 - \mathbb{P}_{(1)}(s,t))I_{(2)}\right]\mathbb{E}\left[p(t)|\mathcal{F}_s\right]. \quad \text{(D.18)}
\end{aligned}$$

Since we in both $\mathbb{P}_{(1)}(s,t)$ and $\mathbb{E}[p(t)|\mathcal{F}_s]$ can allow seasonal time variation, the expected value of $S(t)$ will thus be affected by these specifications. Eq. (D.18) has a very intuitive interpretation: The expected value of day-ahead prices is a weighting of 1) the expected value of day-ahead prices if EPC is present, and 2) the expected domestic prices. The weights are given by the probability of being in each state.

4 Estimation

Due to the assumed independence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$, the estimation procedure can be divided in two independent parts. Namely, estimation of the parameters governing $p(t)$, and estimation of the parameters governing $z(t)$. A common feature is the latent behaviour of the considered processes. We therefore use filtering techniques in both cases.

Besides these two components, we also consider estimation of the weights constituting $I(t)$ inspired by spatial econometrics, which turns out to be estimation of w in our two-market case, cf. Eq. (D.2). We consider each of them in turn in the following sections.

4.1 Estimation of $p(t)$

We assume that we have an equidistant sampling between observations and mark it by Δ. Eqs. (D.7)–(D.8) implies that the state equation of $q(t)$ in a state space representation is

$$q(t) = F(t)q(t-\Delta) + \eta(t), \tag{D.19}$$

where $F(t) = e^{-\Delta K}$, $\eta(t) \sim \mathrm{N}(0, Q_t)$, and $Q_t = \mathbb{V}\mathrm{ar}[q(t)|\mathcal{F}_{t-\Delta}]$. Given the estimation purpose of this section, we introduce a measurement error in Eq. (D.1) and define the measurement equation by

$$\begin{aligned} S(t) &= I(t)p(t) + \epsilon(t) \\ &= I(t)(\Lambda(t) + q(t)) + \epsilon(t), \end{aligned}$$

where $\epsilon(t) \sim \mathrm{N}(0, R(t))$. $R(t)$ is

$$R(t) = \begin{cases} R_{(1)}, & \text{if } I(t) = I_{(1)} \\ R_{(2)}, & \text{if } I(t) = I_{(2)} \end{cases},$$

with $R_{(1)}$ and $R_{(2)}$ being constant diagonal matrices corresponding to the EPC and non-EPC state, respectively.

Kalman Filtering

The normality and linearity of the state space representation lead us to using the Kalman filter in combination with maximum likelihood to estimate the parameters of $p(t)$. There is vast literature on the Kalman filter; see e.g. Schwartz & Smith (2000) in the context of commodity markets.

During the predict step of the Kalman filtering procedure, the a priori estimate of the state is

$$\hat{q}(t|t-\Delta) = F(t)\hat{q}(t-\Delta|t-\Delta),$$

and the a priori estimate of covariance of this state estimate is

$$\hat{P}(t|t-\Delta) = F(t)\hat{P}(t-\Delta|t-\Delta)F(t)^T + Q(t).$$

We update the state estimate and the corresponding covariance by

$$\hat{q}(t|t) = \hat{q}(t|t-\Delta) + H(t)(S(t) - I(t)(\Lambda(t) + \hat{q}(t|t-\Delta))),$$

and

$$\hat{P}(t|t) = \hat{P}(t|t-\Delta) - H(t)I(t)\hat{P}(t|t-\Delta),$$

respectively, where

$$H(t) = \hat{P}(t|t-\Delta)I(t)^T(I(t)\hat{P}(t|t-\Delta)I(t)^T + R(t))^{-1}$$

is the Kalman gain. Due to the assumed independence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$, we regard $I(t)$ as given at time t in the Kalman filtering scheme.

Let T denote the number of observations, and let θ_Λ denote the parameters governing the deterministic function $\Lambda(t)$. Having established the Kalman filtering scheme, we maximize the log-likelihood

$$\mathcal{L}(\theta_q) = \sum_{i=2}^{T} \log f(S(i)|\mathcal{F}_i^{B_z}, \mathcal{F}_{i-1}^{B_p}),$$

with respect to $\theta_q = (\theta_\Lambda, K, \Sigma, R_{(1)}, R_{(2)})$, and where

$$f(S(i)|\mathcal{F}_i^{B_z}, \mathcal{F}_{i-1}^{B_p}) = |2\pi V(i)|^{-1/2} e^{-\frac{1}{2}(S(i)-E(i))^T V(i)^{-1}(S(i)-E(i))},$$

$$E(i) = I(i)(\Lambda(i) + \hat{q}(i|i-1)),$$

and

$$V(i) = I(i)\hat{P}(i|i-1)I(i)^T + R(i).$$

4.2 Estimation of $z(t)$

Let $y(t) \in \{0,1\}$ denote the observed state of $I(t)$ at time t, meaning that $y(t)$ is a Bernoulli process with dependent and non-identically distributed Bernoulli trials. To see this, we assume that $y(t) = 1$ corresponds to EPC. Therefore, by remembering Eq. (D.17) and Prop. 4, we get

$$\begin{aligned}\mathbb{P}(y(t) = 1|\mathcal{F}_s) &= \mathbb{P}(z(t) > 0|\mathcal{F}_s) \\ &= 1 - \Phi\left(\frac{-\mathbb{E}[z(t)|\mathcal{F}_s]}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right) \\ &= 1 - \Phi\left(\frac{-\left(\lambda(t) + e^{-k(t-s)}x(s)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right). \end{aligned} \tag{D.20}$$

Hence, due to the dependence on $x(s)$ we have that the Bernoulli trials are dependent and non-identically distributed.

As a result of Eq. (D.20), the conditional density of $y(t)$ is

$$\begin{aligned} f(y(t)|\mathcal{F}_s) = &\left(1 - \Phi\left(\frac{-\left(\lambda(t) + e^{-k(t-s)}x(s)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right)\right)^{y(t)} \\ &\times \Phi\left(\frac{-\left(\lambda(t) + e^{-k(t-s)}x(s)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right)^{1-y(t)}. \end{aligned} \tag{D.21}$$

Standard estimation methods such as maximum likelihood estimation is not feasible, since the log-likelihood,

$$\begin{aligned} \mathcal{L}(\theta_z) = \sum_{i=2}^{T} \log &\left(1 - \Phi\left(\frac{-\left(\lambda(i) + e^{-k\Delta}x(i-1)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(i)|\mathcal{F}_{i-1}]}}\right)\right)^{y(i)} \\ &\times \Phi\left(\frac{-\left(\lambda(i) + e^{-k\Delta}x(i-1)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(i)|\mathcal{F}_{i-1}]}}\right)^{1-y(i)}, \end{aligned}$$

where $\theta_z = (a_\lambda, b_\lambda, c_\lambda, k, a_\sigma, b_\sigma, c_\sigma)$ cf. Sec. 3.2, depends on the latent process $x(t)$. The linear Kalman filter is suitable in the estimation of $p(t)$, but the non-differentiable relationship between $y(t)$ and $z(t)$ additionally invalidates the extended Kalman filter, and we are thus in the need for another approach.

Particle Filtering

We choose to use particle filtering to mitigate this, and specifically we have employed the bootstrap filter. For references, the interested reader is advised

to look at Kantas *et al.* (2015) in a general state-space model, whereas Pitt *et al.* (2014) consider particle filtering in the context of stochastic volatility models. As in the Kalman filter case, the bootstrap filter can be divided in two steps: a predict and an update step[2]. In our context, the bootstrap procedure is as follows:

Let again T denote the number of observations. For $t = \Delta, 2\Delta, ..., T - \Delta$ we do the following:

- Predict step

  1. Draw n samples

  $$\hat{x}_i(t|t-\Delta) \sim \mathrm{N}(e^{-k\Delta}\hat{x}_i(t-\Delta|t-\Delta), \mathbb{V}\mathrm{ar}[x(t)|\mathcal{F}_{t-\Delta}]),$$

  cf. Prop. 4.

- Update step

  2. For each $i \in \{1, 2, ..., n\}$, calculate weights

  $$\gamma_i(t+\Delta) = \left(1 - \Phi\left(\frac{-\left(\lambda(t+\Delta) + e^{-k\Delta}\hat{x}_i(t|t-\Delta)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(t+\Delta)|\mathcal{F}_t]}}\right)\right)^{y(t+\Delta)} \times \Phi\left(\frac{-\left(\lambda(t+\Delta) + e^{-k\Delta}\hat{x}_i(t|t-\Delta)\right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(t+\Delta)|\mathcal{F}_t]}}\right)^{1-y(t+\Delta)},$$

  cf. Eq. (D.21).

  3. For each $i \in \{1, 2, ..., n\}$, calculate normalized weights,

  $$\tilde{\gamma}_i(t+\Delta) = \frac{\gamma_i(t+\Delta)}{\sum_{j=1}^{M}\gamma_j(t+\Delta)}.$$

  4. For each $i \in \{1, 2, ..., n\}$, draw a sample $\hat{x}_i(t|t)$ from the distribution with density

  $$f(x) = \sum_{j=1}^{n}\tilde{\gamma}_j(t+\Delta)\delta(x - \hat{x}_j(t|t-\Delta)),$$

  where $\delta(\cdot)$ is Dirac's delta function.

[2]The bootstrap filter can also be divided in an importance and selection step, but we find it more intuitive to divide the algorithm in a predict and update step, also making the Kalman filter presented earlier and the bootstrap filter easily comparable.

Having obtained n values of $x(t)$ (i.e., $z(t)$) for each $t = 0, \Delta, ..., T - \Delta$, we evaluate the log-likelihood as

$$\mathcal{L}(\theta_z) = \sum_{i=2}^{T} \log \left(\frac{1}{n} \sum_{j=1}^{n} \left(1 - \Phi \left(\frac{-\left(\lambda(i) + e^{-k\Delta} \hat{x}_j(i-1|i-2) \right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(i)|\mathcal{F}_{i-1}]}} \right) \right)^{y(i)} \times \Phi \left(\frac{-\left(\lambda(i) + e^{-k\Delta} \hat{x}_j(i-1|i-2) \right)}{\sqrt{\mathbb{V}\mathrm{ar}[z(i)|\mathcal{F}_{i-1}]}} \right)^{1-y(i)} \right).$$

As pointed out by Pitt *et al.* (2014), one single run of the particle filter will therefore estimate the log-likelihood.

Due to the non-observable nature of the $z(t)$ process and the fact that the sign of it only matters, one has to fix the level of it in order to avoid possible oscillating estimation issues. Hence, we fix $\sigma(t)$ to oscillate around 1, by setting a_σ to

$$a_\sigma = \sqrt{\frac{2k}{1 - e^{-2k\Delta}}}.$$

4.3 Estimation of $I(t)$

Inspired by weight matrix determination in spatial econometrics (see e.g. Abate & Haldrup (2017)), we construct $I(t)$ using the spatial relationship between the market areas. Although our model differs from spatial econometric models in that the observed prices do not depend on lagged values of observed prices from the market areas under consideration, but rather depend on a weighting of (partly) unobservable domestic prices, the weighting of the unobservable domestic prices are chosen in a similar way. As discussed in Abate & Haldrup (2017), choosing the spatial weight matrix is a delicate task, due to the selection of the factor influencing the weight matrix. Since net export basically results in a horizontal shift in either the domestic supply or demand curve, depending on the sign of the net export cf. Fig. D.1, the consumed electricity in each market area determines how sensitive the resulting domestic market price is to variation in net export. Loosely speaking, a large market area—based on the supplied electricity—is not as price sensitive to net export as a smaller market area. Hence, we use the supplied electricity in a market area as the "size" of the market area. In the two-market case, estimation of $I(t)$ boils down to estimating w, cf. Eq. (D.2).

5 Application

In this section, we first exemplify the proposed model by jointly modelling hour 21 day-ahead prices for Germany and France. Presentation of estimation results is then followed by an analysis of derivative prices implied by the model.

5.1 Case Study of German and French Electricity Prices

Our data set introduced in Sec. 2 contains daily observations. Hence, we set $M = 365$, and the time between adjacent observations is set to $1/365$. We report estimation results for $p(t)$ and $z(t)$ in the following sections after a short note on the estimation of w in $I_{(1)}$. As discussed in Sec. 4.3, we use the supplied electricity to determine w. Table D.2 reports the supplied

DE	FR
575,701	496,932

Table D.2: Supplied electricity given in (GWh) based on IEA (2017).

electricity for Germany and France, which results in a weight for Germany of $575,701/(575,701 + 496,932) = 0.537$. The resulting two states of $I(t)$ are given in Eq. (D.22).

$$I_{(1)} = \begin{bmatrix} 0.537 & 0.463 \\ 0.537 & 0.463 \end{bmatrix} \text{ or } I_{(2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{D.22}$$

Motivated by the empirical findings in Sec. 2 and related literature on electricity price modeling (e.g. Pircalabu & Benth (2017)), we define the seasonal function $\Lambda(t)$ in Eq. (D.3) as

$$\Lambda(t) = a_\Lambda + b_\Lambda \cos\left(\frac{2\pi t}{M}\right) + c_\Lambda \sin\left(\frac{2\pi t}{M}\right) + d_\Lambda I_W(t),$$

where $I_W(t)$ is an indicator function being one if day t is either Saturday or Sunday, and $a_\Lambda, b_\Lambda, c_\Lambda, d_\Lambda \in \mathbb{R}^2$ are constants. We further assume that Σ is lower triangular, and that K is diagonal in Eq. (D.4). In Tables D.3-D.4 the estimation results for θ_q are reported using the estimation method presented in Sec. 4. To simplify the presentation, the estimation results are divided in the parameters governing $\Lambda(t)$ and $q(t)$.

The estimated weekend coefficient, $\hat{d_\Lambda}$ for both market areas is negative, and thus reflects the lower electricity demand during weekends compared to weekdays. Also notable is the estimated base level $\hat{a}_\Lambda$, where the German base level is lower than the French base level. Regarding the speed of mean

	$\hat{a}_\Lambda$	$\hat{b}_\Lambda$	$\hat{c}_\Lambda$	$\hat{d}_\Lambda$
DE	36.42	0.75	3.43	-4.74
FR	45.17	4.52	-7.65	-6.34

Table D.3: Estimation results for $\Lambda(t)$.

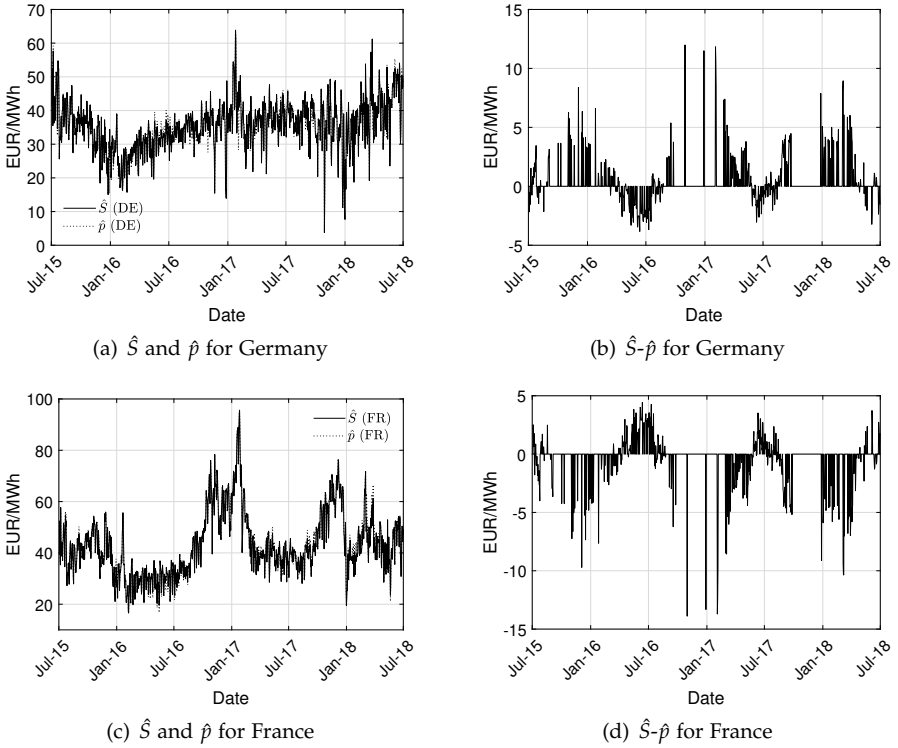
$\hat{K}_{(1,1)}$	$\hat{K}_{(2,2)}$	$\hat{\Sigma}_{(1,1)}$	$\hat{\Sigma}_{(2,1)}$	$\hat{\Sigma}_{(2,2)}$
89.53	30.68	99.95	52.34	52.15

Table D.4: Estimation results for $q(t)$.

reversion, the estimated German one, $\hat{K}_{(1,1)}$, corresponds to a half-life of approximately 2.8 days, whereas the estimated French speed of mean reversion corresponds to a more persistent behaviour with a half-life of approximately 8.2 days.

To investigate the estimation results in more detail, we obtain the estimated $\hat{p}$ process by filtering of $\hat{q}$, following the Kalman scheme described in Sec. 4. In addition, using the estimated forms of $I(t)$ presented in Eq. (D.22), and assuming we have knowledge of $I(t)$ at time t as in the described Kalman scheme, we also obtain the estimated $\hat{S}$ process. The intention of showing the estimated $\hat{S}$ process in addition to the estimated $\hat{p}$ process is twofold. Firstly, to give a sense of the model's ability to replicate the observed prices in Fig. D.3(a), and secondly to give an insight in the difference between the model implied $\hat{S}$ and $\hat{p}$ processes. In Figs. D.6(a) and D.6(c) we show the estimated $\hat{p}$ process and the corresponding estimated $\hat{S}$ process for both Germany and France, resembling the observations in Fig. D.3(a) quite well. Figs. D.6(b) and D.6(d) illustrate the difference between the estimated $\hat{S}$ and $\hat{p}$ processes. It is apparent that the differences are more frequent during spring/summer than autumn/winter, reflecting the fact that EPC happens more frequently during spring/summer. Also notable is the fact that the differences are greater during autumn/winter, so from this perspective, the impact from EPC on a single observation is highest during autumn/winter. Due to the parameter w affecting the weighting of the domestic day-ahead prices (the p process), the resulting differences between the $\hat{S}$ and $\hat{p}$ processes in the two market areas differ. The pattern in Fig. D.6(b) is therefore reversed in Fig. D.6(d), but they differ in absolute terms.

The estimated parameters for the $z(t)$ process are reported in Table D.5. The estimated speed of mean reversion translates to a half-life of approximately 2.56 days. Fig. D.7(a) shows the estimated asymptotic probability of observing EPC throughout the year. The yearly pattern observed in Fig. D.5(a) is thus replicated by the model. The significant drop in probabil-



(a) $\hat{S}$ and $\hat{p}$ for Germany

(b) $\hat{S}$-$\hat{p}$ for Germany

(c) $\hat{S}$ and $\hat{p}$ for France

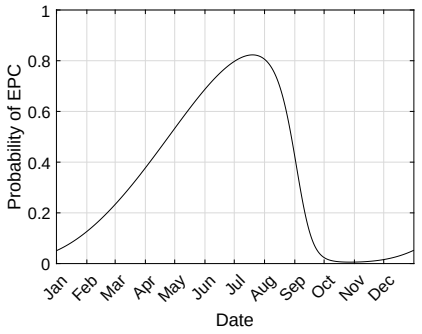
(d) $\hat{S}$-$\hat{p}$ for France

Fig. D.6: (a) Estimated $\hat{S}$ and $\hat{p}$ for Germany, (c) $\hat{S}$-$\hat{p}$ for Germany, (c) estimated $\hat{S}$ and $\hat{p}$ for France, (d) $\hat{S}$-$\hat{p}$ for France.

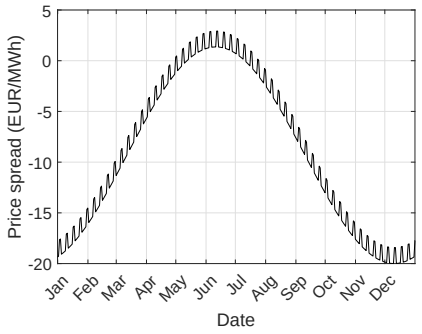
$\hat{k}$	$\hat{a}_\lambda$	$\hat{b}_\lambda$	$\hat{c}_\lambda$	$\hat{b}_\sigma$	$\hat{c}_\sigma$
98.8570	-0.6146	1.3094	0.0565	-14.2477	-1.3537

Table D.5: Estimation results for $z(t)$.

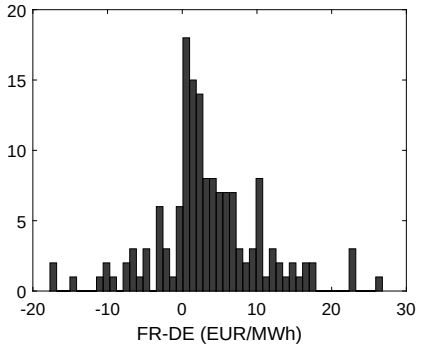
ity during August/September is notable and marks the switch from summer to autumn. The lowest probability is reached in November and is as low as 0.2%, while the highest is in July and is 82%. In a risk management or asset valuation situation, this is quite a difference and must be taken into account. We elaborate on this matter when covering derivative prices implied by the model.



(a) Asymptotic probability of EPC.



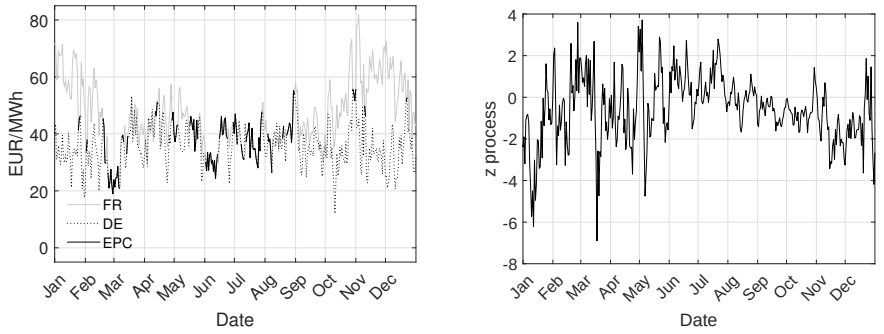
(b) Difference between the seasonality functions.



(c) FR-DE day-ahead price spread.

Fig. D.7: (a) Estimated asymptotic probability of observing EPC throughout a year, (b) difference between the estimated seasonality functions affecting the p process, and (c) the observed German hour 21 day-ahead price subtracted from the French hour 21 day-ahead price. The prices used in (c) are prices from non-EPC days, where we observe EPC the next day.

Remember that we assume independence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$, which implies that we can divide the estimation procedure in two independent parts. Fig. D.7(c) shows a histogram of the observed price spread between France and Germany. To investigate the dependence between the $p(t)$ process and the $z(t)$ process, we only show price spreads from non-EPC days, which is immediately followed by days with EPC. Most occurrences are observed just about zero, meaning that there is a tendency that EPC hap-



(a) Simulated DE and FR prices (b) Simulated z process

Fig. D.8: (a) Simulated hour 21 day-ahead prices for Germany and France, and (b) corresponding simulation of z process.

pens in periods with small price spreads. Though, there is no guarantee that EPC only happens in these periods, since it reaches as high as approximately 27 EUR/MWh in absolute terms. From the perspective of the estimated model, Fig. D.7(b) shows the difference between the asymptotic mean levels of $p(t)$; i.e., it shows the spread between $\Lambda_1(t)$ and $\Lambda_2(t)$, where $\Lambda(t) = [\Lambda_1(t) \quad \Lambda_2(t)]^\top$. Comparing the shape of Fig. D.7(b) to the shape of Fig. D.7(a), we conclude that a high probability of EPC roughly corresponds to a small price spread asymptotically, and vice versa. Loosely speaking, even though we assume independence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$, the seasonal components of the model suggest that a high probability of EPC translates to a small spread between the components of $p(t)$, and that a low probability of EPC translates to a large spread between the components of $p(t)$. Assuming independence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$ thus has the advantage that we can split up the estimation procedure in two independent parts, while still allowing the above discussed dependency between the EPC probability and the magnitude of the spread between the components of $p(t)$. We leave it as future work to investigate the case with dependence between $\mathcal{F}_t^{B_z}$ and $\mathcal{F}_t^{B_p}$.

To further illustrate the model, we show in Fig. D.8(a) a simulated path for both the German and French prices using the parameters reported in Tables D.3-D.5. Again we emphasize when EPC occurs by a solid black line. As expected from Fig. D.7(a) we observe most EPC states during summer, and least during autumn/winter. Regarding the EPC aspect of the model, we plot in Fig. D.8(b) the simulated $z(t)$ process, which is, due to the construction, positive whenever we observe EPC in Fig. D.8(a).

5.2 Derivative Pricing

In electricity markets, the underlying cannot be stored—at least not effectively—and the classical pricing assumptions about market completeness do therefore not hold. Hence, one typically resorts to imposing a parametric form through a measure change, and thereby introducing the so-called market price of risk. By using market data, it is then possible to back out the market price of risk.

The Girsanov theorem gives the p-process under an equivalent pricing measure Q as

$$p(t) = \Lambda(t) + e^{-(t-s)K}q(s) + \int_s^t e^{-(t-u)K}\Sigma\varphi(u)du + \int_s^t e^{-(t-u)K}\Sigma dB_p^Q(u). \tag{D.23}$$

where $\varphi(t) \in \mathbb{R}^N$ is the so-called market price of risk, assumed to be adapted to $\{\mathcal{F}_t\}_{t\geq 0}$ and satisfying certain integrability conditions (see Girsanov (1960)). In addition, $B_p^Q(t) \in \mathbb{R}^N$ is a Brownian motion under the new measure Q. By assuming no market price of risk concerning $z(t)$, we thus see that an appropriate shift in the mean level of $p(t)$ will yield the appropriate measure change.

In practice, one might assume that $\varphi(t)$ is a piecewise constant deterministic function and match it to the observed forward prices at time s, which we will cover in the next section. Eq. (D.23) then implies that $p(t)$ under Q is normally distributed with expected value

$$\mathbb{E}^Q[p(t)|\mathcal{F}_s] = \Lambda(t) + e^{-(t-s)K}q(s) + \int_s^t e^{-(t-u)K}\Sigma\varphi(u)du \tag{D.24}$$

and variance given in Eq. (D.6). It follows from Eq. (D.24) and Prop. 5 that $S(t)$ under Q is characterized by the density

$$f_{s,t}^Q(x) = \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t)\frac{1}{(2\pi)^{N_i/2}|\Sigma_{(i)}(s,t)|}e^{-\frac{1}{2}\left(x-\mu_{(i)}^Q(s,t)\right)^{\mathrm{T}}\Sigma_{(i)}(s,t)^{-1}\left(x-\mu_{(i)}^Q(s,t)\right)}, \tag{D.25}$$

where

$$\mu_{(i)}^Q(s,t) = I_{(i)}\mathbb{E}^Q[p(t)|\mathcal{F}_s], \tag{D.26}$$

and everything else is defined in Prop. 5.

Forwards

The liquidity of electricity markets is low compared to traditional financial markets, with the forward market in general being the most liquid. As a

result, derivative prices are often restricted to be consistent with the forward market[3], which is used to back out the market price of risk. Following the lines of the discussion of market incompleteness in electricity markets, we *define* the forward price at time s for a forward contract with delivery time t as the expected value under Q of the future day-ahead price, as in e.g. Benth & Meyer-Brandis (2009). In terms of our multivariate model, we have

$$F(s,t) = \mathbb{E}^Q[S(t)|\mathcal{F}_s],$$

where $F(s,t) \in \mathbb{R}^N$ is the forward price at time s for delivery at time t for N different market areas. In our modeling framework, Prop. 6 states the forward price.

Proposition 6
Let $S(t)$ be defined under Q as in Eq. (D.25). Then the forward price at time s for delivery at time t for two different market areas is

$$F(s,t) = \left[\mathbb{P}_{(1)}(s,t)I_{(1)} + \mathbb{P}_{(2)}(s,t)I_{(2)}\right] \mathbb{E}^Q\left[p(t)|\mathcal{F}_s\right]. \tag{D.27}$$

Proof. Follows from Eq. (D.18) and Eq. (D.25). □

Eq. (D.27) implies that the spread between forward prices decreases gradually as $\mathbb{P}_{(1)}(s,t) \to 1$. In the extreme case where $\mathbb{P}_{(1)}(s,t) = 1$, the model suggest that the forward prices in the two different market areas are equal. The model is hence in line with the expectation of a single forward price in a situation with two perfectly coupled markets.

Transmission Rights

A transmission right is a European spread option with the underlying assets being day-ahead prices of adjacent market areas. The pricing formula for transmission rights given the model presented in Sec. 3 is closely related to the work by Bachelier (1900) and is given in the next proposition.

Proposition 7
Let $S(t)$ be defined under Q as in Eq. (D.25), $g \in \mathbb{R}^2$, and let $(g^{\mathrm{T}}s(t) - H)^+$ be the payoff of a European spread option with strike H. Then

$$\begin{aligned}\mathbb{E}^Q[(g^{\mathrm{T}}S(t) - H)^+|\mathcal{F}_s] = \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t) \Bigg(& a_{(i)}(s,t)\Phi\left(\frac{a_{(i)}(s,t)}{c_{(i)}(s,t)}\right) \\ & + c_{(i)}(s,t)\phi\left(\frac{a_{(i)}(s,t)}{c_{(i)}(s,t)}\right)\Bigg),\end{aligned} \tag{D.28}$$

[3]We assume a deterministic interest rate, and thus the term forward market does also cover the futures market.

where

$$a_i(s,t) = g^{\mathrm{T}}\mu^{Q}_{(i)}(s,t) - H, \tag{D.29}$$

$$c_i(s,t) = \sqrt{g^{\mathrm{T}}\Sigma_{(i)}(s,t)g}, \tag{D.30}$$

and Φ and ϕ is the distribution and density function of the standard normal distribution, respectively. $\mu^{Q}_{(i)}(s,t)$ is given in Eq. (D.26), and $\Sigma_{(i)}(s,t)$ is given in Prop. 5.

Proof. Define $Z(t) := g^{\mathrm{T}}S(t) - H$. Eq. (D.25) implies that $Z(t)$ under Q is

$$Z(t)|\mathcal{F}_s \sim \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t)N(g^{\mathrm{T}}\mu^{Q}_{(i)}(s,t) - H, g^{\mathrm{T}}\Sigma_{(i)}(s,t)g)$$

Hence, by defining $a_i(s,t)$ and $c_i(s,t)$ as in Eqs. (D.29)-(D.30) and $U \sim \mathrm{N}(0,1)$, we obtain

$$\begin{aligned}
\mathbb{E}^{Q}[Z(t)^{+}|\mathcal{F}_s] &= \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t)\mathbb{E}^{Q}[(a_i(s,t) + c_i(s,t)U)^{+}|\mathcal{F}_s] \\
&= \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t) \int_{-a_i(s,t)/c_i(s,t)}^{\infty} \frac{a_i(s,t) + c_i(s,t)u}{\sqrt{2\pi}} e^{-u^2/2} du \\
&= \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t) \left(a_i(s,t)\Phi\left(\frac{a_i(s,t)}{c_i(s,t)}\right) + c_i(s,t)\phi\left(\frac{a_i(s,t)}{c_i(s,t)}\right)\right),
\end{aligned}$$

where $\mathbb{P}_{(i)}(s,t)$ is defined in Eq. (D.16). □

To further comment on Prop. 7, we set $g = [1 \quad -1]$, and $H = 0$, which is also the most relevant case having the transmission rights in Europe in mind. Consequently, we get that $g^{\mathrm{T}}S(t) - H = 0$ in the EPC case, and Eq. (D.28) thus reduces to

$$\begin{aligned}
\mathbb{E}^{Q}[(S_1(t) - S_2(t))^{+}|\mathcal{F}_s] =& \mathbb{P}_{(2)}(s,t)\left(a_{(2)}(s,t)\Phi\left(\frac{a_{(2)}(s,t)}{c_{(2)}(s,t)}\right)\right. \\
&\left. + c_{(2)}(s,t)\phi\left(\frac{a_{(2)}(s,t)}{c_{(2)}(s,t)}\right)\right).
\end{aligned} \tag{D.31}$$

Intuitively, the transmission right value should be increasing when the probability of no convergence, $\mathbb{P}_{(2)}(s,t)$, increases; this is confirmed by Eq. (D.31). As a result, by assuming no price convergence—i.e. $\mathbb{P}_{(2)}(s,t) = 1$—one would simply overestimate the value of the transmission right. Also, having Fig. D.7(a) in mind and the empirical setup in Sec. 5.1, there is a significant difference between the reduction of transmission right prices throughout the year, simply due to the time-varying probability of EPC.

6 Conclusion

In this paper, we propose a model for the joint behaviour of day-ahead electricity prices in a two-market setup where the market coupling mechanism is taken into account in a continuous time manner. The model is dividable in two parts inspired by the way the market coupling mechanism is effectuated in Europe: The first part captures the exact price convergence, modeled by a latent univariate Gaussian Ornstein-Uhlenbeck process. The second part can be interpreted as capturing the domestic prices set in each individual market area, which we model by a partly latent multivariate Gaussian Ornstein-Uhlenbeck process.

We derive theoretical results of the model, and we discuss the estimation of the model in detail. With regard to the latter, our estimation proposal consists of two assumed independent parts: 1) standard Kalman filtering to estimate the process capturing the domestic prices, and 2) particle filtering to estimate the process capturing the exact price convergence. Both approaches are used in combination with maximum likelihood.

The model is exemplified through a case study where we jointly model German and French hour 21 day-ahead prices. We find a strong seasonal behaviour in the estimated probability of observing exact price convergence. Lastly, we discuss derivative prices implied by the model and show that the implied forward and transmission right prices can be obtained by closed-form formulas. The impact of the exact price convergence on forward and transmission right prices can be directly read off from the closed-form formulas. In the example with German and French hour 21 day-ahead prices, the strong seasonality in the probability of observing exact price convergence suggests that 1) the spread between forward prices throughout a year varies greatly, and 2) with everything else being equal the transmission right prices throughout a year vary greatly.

Future research includes the case where the number of market areas exceeds two. Although we have argued that such case is not relevant given the presently traded instruments, advanced scenario analysis could be derived from such model. Furthermore, it would be interesting to investigate how to incorporate, and the implication of including, spikes. As an example, large movements in opposite directions of two prices would significantly impact the value of a transmission right. However, as a consequence of the market coupling mechanism, such scenarios are very rare, and spikes of the same sign are often observed in adjacent market areas. The real impact of such spikes on e.g. derivatives could therefore be of minor importance. This is also supported by the fact that fewer and fewer spikes are observed, which in turn is possible due to the increased market interconnectedness that potentially eliminates spikes.

Acknowledgements

The authors are grateful for the valuable comments and discussions provided by the participants attending the Quantitative Methods in Finance 2018 Conference in Sydney. The authors would also like to thank one anonymous reviewer for helpful comments that improved the presentation of this paper.

Disclosure Statement

No conflict of interest to declare.

Funding

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B.

References

Abate, G. D., Haldrup, N. (2017). Space-time modeling of electricity spot prices. *The Energy Journal*, **38**, pp. 175–196.

Bachelier, L. (1900). Théorie de la spéculation. *Ann. Sci. Écol. Normale Supér.*, **17**, pp. 21–86.

Benth, F. E., Šaltytė Benth, J. (2004). The normal inverse Gaussian distribution and spot price modelling in energy markets. *International Journal of Theoretical and Applied Finance*, **07**, pp. 177–192.

Benth, F. E., Kallsen, J., Meyer-Brandis, T. (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, **14**, pp. 153–169.

Benth, F. E., Meyer-Brandis, T. (2009). The information premium for non-storable commodities. *Journal of Energy Markets*, **2**, 111–140.

Borovkova, S., Schmeck, M. D. (2017). Electricity price modeling with stochastic time change. *Energy Economics*, **63**, pp. 51–65.

Douglas, S. M., Popova, J. N. (2011). Econometric Estimation of Spatial Patterns in Electricity Prices. *The Energy Journal*, **32**, pp. 81—105.

Epex Spot, A major step towards market integration, 2019. `http://www.epexspot.com/en/market-coupling` (accessed 11 July 2019).

Elia, CWE flow-based, 2019. `http://www.elia.be/en/products-and-services/cross-border-mechanisms/transmission-capacity-at-borders/flow-based-marktkoppeling-centr-w-europa` (accessed 11 July 2019).

Geman, H., Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *Journal of Business*, **79**, pp. 1225–1261.

Girsanov, I. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, **5**, pp. 285–301.

Haldrup, N., Nielsen, F. S., Nielsen, M. Ø. (2010). A vector autoregressive model for electricity prices subject to long memory and regime switching. *Energy Economics*, **32**, pp. 1044–1058.

Hinderks, W., Wagner, A., Korn, R. (2018). A structural Heath-Jarrow-Morton framework for consistent intraday, spot, and futures electricity prices. *arXiv*.

Hubbard., R. A., Inoue, L. Y. T., Fann, J. R. (2014). Modeling nonhomogeneous Markov processes via time transformation. *Biometrics*, **64**, pp. 843–850.

Ichihara, K., Kunita, H. (1974). A classification of the second order elliptic operator and its probabilistic characterization. *Z. Wahrsch. Verv. Geb.*, **30**, pp. 235–254.

Ignatieva, K., Trück, S. (2016). Modeling spot price dependence in Australian electricity markets with applications to risk management. *Computers & Operations Research*, **66**, pp. 415–433.

International Energy Agency (2017). Monthly electricity statistics. December 2017 report for electricity production and trade data for all OECD member countries.

Inamura, Y. (2006). Estimating continuous time transition matrices from discretely observed data. *Bank of Japan Working Paper Series.*.

Kantas, N., Doucet, A., Singh, S. S., Maciejowski, J., Chopin, N. (2015). On particle methods for parameter estimation in state-space models. *Statistical Science*, **30**, pp. 328–351.

Laub., A. J. (2005). Matrix Analysis for Scientists and Engineers. *Society for Industrial and Applied Mathematics*, ISBN: 978-0-898715-76-7.

Lucia, J. J., Schwartz, E. (2002). Electricity prices and power derivatives: evidence from the Nordic power exchange. *Review of Derivatives Research*, **5**, pp. 5–50.

Macedo, H. D., Oliveira, N. O. (2013). Typing linear algebra: A biproduct-oriented approach. *Science of Computer Programming*, **78**, pp. 2160–2191.

Mahringer, S., Füss, R., Prokopczuk, M. (2015). Electricity market coupling and the pricing of transmission rights: an option-based approach. *Working papers on Finance 1512. University of St. Gallen, School of Finance*.

Meyer-Brandis, T., Tankov, P. (2008). Multi-factor jump-diffusion models of electricity prices. *International Journal of Theoretical and Applied Finance*, **11**, pp. 503–528.

Park, H., Mjelde, J. W., Bessler, D. A. (2006). Price dynamics among U.S. electricity spot markets. *Energy Economics*, **28**, pp. 81–101.

Pircalabu, A., Benth, F. E. (2017). A regime-switching copula approach to modeling day-ahead prices in coupled electricity markets. *Energy Economics*, **68**, pp. 283–302.

Pitt, M. K., Malik, S., Doucet, A. (2014). Simulated likelihood inference for stochastic volatility models using continuous particle filtering. *Annals of the Institute of Statistical Mathematics*, **66**, pp. 527–552.

Schwartz, E., Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, **46**, pp. 893–911.

Paper E

Trading in Coupled Interconnected Electricity Markets: A Transmission Right Perspective

List of authors: Troels Sønderby Christensen[1,2], Jesper Jung[2]

[1]Department of Mathematical Sciences, Aalborg University
[2]Quantitative Analytics, Centrica Energy Trading

The paper has been submitted to Energy Economics.

The layout has been revised.

Abstract

Transmission rights facilitate the handling of the risk associated by trading electricity in adjacent market areas. With the increased interconnectedness in European electricity markets, the amount of time that one observes the special behaviour of zero price spreads between coupled markets has increased. In this paper, we investigate pricing and hedging of transmission rights in coupled electricity markets using the recently proposed Conv model from Christensen & Benth (2019) and the well-known option spread formula from Margrabe (1978) as a benchmark. We discuss the properties of the Conv model and the Margrabe formula in a transmission right context, followed by an empirical investigation. We find that the Conv model which takes into account the market coupling significantly outperforms the Margrabe formula in a pricing context. In a hedging context the two models perform equally well on the data we have used for our analysis.

1 Introduction

Interconnected electricity markets is an important cornerstone to benefit significantly from the increasing non-programmable renewable electricity sources such as wind and solar power production. A surplus of cheap wind power production in one market area should in a proper interconnected market be transferred to other market areas to the benefit of both the exporting and importing parties. In Europe, the liberalization of electricity markets has incentivized such actions. The overall goal from the European Commission is a single pan-European market, where electricity can flow uncongested, ultimately leading to one single European electricity price [see Füss *et al.* (2015)]. Historically, ex-ante auctions for daily rights to transfer electricity from one market area to another was applied. Ex-ante refers to the fact that market participants had to schedule the flow of electricity for their transmission rights before actually knowing the day-ahead electricity prices. Hence, market participants could schedule flows inefficiently, potentially moving electricity from a high price area to a low price area. Note that day-ahead prices denote the price of electricity for each hour of the following day, settled one day prior to delivery.

To mitigate inefficient scheduling, implicit auctions for the daily transmission rights, also known as market coupling, have been introduced in European electricity markets. The market coupling mechanism starts with market participants submitting their bids and offers for electricity in the interconnected market areas. One single algorithm then calculates optimal electricity flows and the corresponding day-ahead electricity prices to optimize the overall social welfare. Historically, the trilateral market coupling was introduced in 2006 where the French, Belgian, and Dutch day-ahead electricity markets

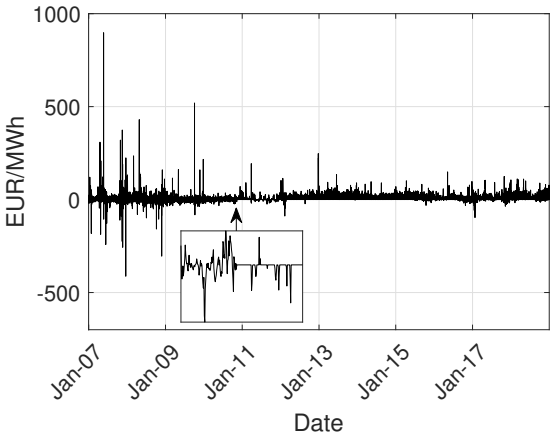


Fig. E.1: Price spread between the Netherlands and Germany for the period from 1 January 2007 to 31 December 2018. The arrow indicates the shift to market coupling between the Netherlands and Germany in November 2010. The zoomed graph emphasizes the radical change in the spread between the Netherlands and Germany before and after November 2010.

were coupled. Germany joined in November 2010, and prior to May 2015, the determination of electricity flows were determined by the so-called available transfer capacity method. In May 2015, this method was replaced by the flow-based method. According to Amprion (2018), the advantage of the flow-based compared to the available transfer capacity method is that it *leads in most cases to a larger flow based domain and higher welfare gains.* The domain refers to the feasible flows of electricity between the market areas. The interested reader is referred to den Bergh *et al.* (2015) for more information on the switch from the available transfer capacity method to the flow-based one.

To exemplify the implication of the market coupling on day-ahead prices, we plot in Fig. E.1 the price spread between hourly day-ahead electricity prices for the Netherlands and Germany (German price subtracted from the Dutch price) for the 12-years period starting from 1 January 2007 and ending at 31 December 2018. As a first observation, one might notice that both the frequency and magnitude of the extreme price spreads seem to fade out, indicating that the increased interconnectedness also serves as a reserve against large price spread movements. Secondly, after November 2010, where Germany was included in the market coupling, the price spread often equals exactly zero, so-called exact price convergence (EPC). Before November 2010, the spread fluctuated around zero and did very rarely equal zero exactly.

Even though EPC happens quite often, there is still a long way to fully uncongested flows of electricity. Fig. E.2 shows a stacked bar plot calculated on a monthly basis for the same period as in Fig. E.1. It shows the percentage of time during a particular month that two or more markets couple

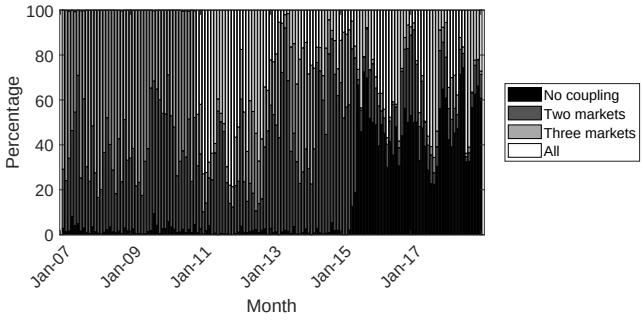


Fig. E.2: Time spent in the CWE area in one of the four states 1) no coupling, 2) two markets couple, 3) three markets couple, and 4) four markets couple. Each month during the period from 1 January 2007 to 31 December 2018 is shown. The label referring to two markets does also include the case where two pairwise markets couple; i.e., Germany and France couple while Belgium and the Netherlands also couple.

in the Central-West European (CWE) price area. Until November 2010, the French, Belgian, and Dutch day-ahead markets coupled most of the time; either all three or just two of them. With Germany joining in November 2010, the dynamic changed and the amount of time that all four markets coupled seemed to replace the amount of time that the French, Belgian, and Dutch day-ahead markets coupled previously. Most notable however is the counter-intuitive fact that the amount of time that no markets couple increases significantly with the implementation of the flow-based market coupling. The increased flexibility in the flows of electricity does apparently not lead to a higher amount of time that the markets couple. Continuing on this line of perspective, although Fig. E.2 shows a decrease in EPC, the level of convergence of the prices might be increasing; i.e., the price spreads between the market areas could have shrunken while the EPC percentages could have shrunken as well. Fig. E.3 shows the mean of the absolute differences between the four price pairs in the CWE area, however, it does not indicate that this is the case. We do in fact see an increase in the minimum level of the absolute price differences for all price pairs.

In this paper we do not investigate further the level of interconnectedness of electricity markets, but the naive analysis just performed is in fact in agreement with more thorough investigations. Here we mention de Menezes & Houllier (2016) who reject the hypothesis that *EU electricity markets are increasingly integrated* in the day-ahead markets. Without pursuing this matter any further we can nevertheless conclude that the special joint behaviour of the day-ahead electricity prices in the CWE area exemplified in Fig. E.1 (and Europe as a whole) seems to continue. Market participants trading electricity in these coupled markets do therefore face the uncertainty of the price spread, and will likely have to do it for a long time. In Europe, transmission

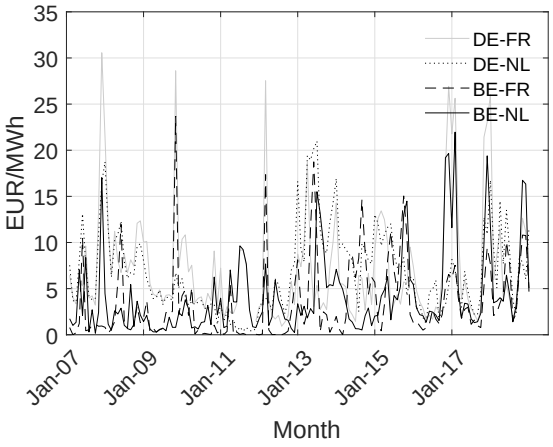


Fig. E.3: Monthly mean of the absolute differences between the four price pairs corresponding to the physical interconnectedness in the CWE area for the period from 1 January 2007 to 31 December 2018.

rights for a longer period than a day, starting from a month, is tradeable. The underlying of such a transmission right is the day-ahead prices, thus allowing market participants to hedge day-ahead price differences between market areas. A transmission right is basically a basket of spread options between two market areas, and can either be physically or financially settled. A physical transmission right will however be equivalent to a financial transmission right, if the former is sold under the widely used "use-it-or-sell-it" principle (see Alasseur & Féron (2018) for more information). We will therefore neglect the distinction between the two and simply use the term transmission right going forward.

The daily average of the hourly day-ahead prices is often used in mark-to-market and risk management contexts, hence the related literature on modeling the average day-ahead price is extensive, see e.g. Lucia & Schwartz (2002) and Geman & Roncoroni (2006), or more recently the work of Borovkova & Schmeck (2017). However, the proportion of time that EPC happens is lost when averaging the hourly prices. For the period from 1 July 2015 to 30 June 2019, the EPC percentages for the four cross-border connections in the Central-West Europe (CWE) area are shown in Table E.1, which clearly shows huge differences in daily and hourly EPC percentages. The literature on modeling day-ahead prices with hourly granularity is not as comprehensive as the literature on modeling the average day-ahead price. Huisman *et al.* (2007) proposed a model for a single market area that takes the panel data feature into account in the sense that all hourly day-ahead prices are set the same time the day prior to delivery. The consideration of more than one market in a market coupling environment is the theme of Pircalabu & Benth

	DE-FR	FR-BE	BE-NL	NL-DE
Daily	2.19	2.67	3.15	2.94
Hourly	35.69	41.97	41.32	40.68

Table E.1: Observed EPC percentage for the period from 1 July 2015 to 30 June 2019 for four different price pairs in the CWE area.

(2017) and Christensen & Benth (2019). In contrast to Huisman *et al.* (2007), each hour is modelled individually. Obviously, handling derivatives written on a combination of at least two different hourly day-ahead prices is not possible using Pircalabu & Benth (2017) and Christensen & Benth (2019) (e.g. a spread option between two consecutive hours for a single market area). Though, these models are perfectly suited for the valuation and hedging of transmission rights, where it is possible to separate the payoff structure on an hourly basis. While the model of Pircalabu & Benth (2017) allow for advanced dependence structures through flexible copula modeling, the model of Christensen & Benth (2019) allows for closed-form formulas for derivatives such as transmission rights. The mentioned articles all have a reduced-form modeling perspective. Structural approaches, where one typically models the offer and bid curves explicitly, have also been investigated in the literature, see Kiesel & Kustermann (2015), Mahringer *et al.* (2015), and Alasseur & Féron (2018). Kiesel & Kustermann (2015) derive an analytical formula for forward prices, while Mahringer *et al.* (2015) and Alasseur & Féron (2018) additionally derive analytical formulas for transmission rights.

In this paper, we investigate the pricing of transmission rights and hedging of these in a market coupling environment. We rely on the model proposed by Christensen & Benth (2019). This model will for the remaining part of the paper be denoted by the Conv model, since it takes the EPC aspect into account. As a benchmark throughout the paper, we also present results using the formula proposed by Margrabe (1978) due to its popularity and well-known way of pricing and hedging spread options. The two models rely on different assumptions. Shortly put, the Margrabe formula relies on the assumption that the underlying prices can be described by geometric Brownian motions, whereas the Conv model takes the market coupling into account, but still allows closed-form formula for spread options and the corresponding greeks.

The remaining part of this paper is organised as follows. In Sec. 2 we introduce transmission rights and considerations of hedging these. In Sec. 3 we introduce and discuss the models used to investigate the pricing and hedging of the transmission rights, which is the theme of Sec. 4. Sec. 5 concludes.

2 Transmission Rights

Let $S(t,h) \in \mathbb{R}^2$ denote the day-ahead electricity price at day t and hour h at two market areas. A transmission right is essentially a basket of European spread options, all with strike zero. Denoting the price in EUR/MWh, the payoff of a transmission right is

$$\frac{1}{\sum_{t=1}^{T} H(t)} \sum_{t=1}^{T} \sum_{h=1}^{H(t)} (g^\top S(t,h))^+, \tag{E.1}$$

where $g \in \mathbb{R}^2$, T is the number of days of the delivery period of the transmission right, and $H(t)$ is the number of hours on day t. Typically, $g = [1 \quad -1]^\top$. As a consequence of the payoff structure, valuing each option, $(g^\top S(t,h))^+$, independently of each other is possible. To ease the presentation we denote the day-ahead electricity price by $S(t)$ going forward, even though we only consider hourly day-ahead prices.

Due to the limited liquidity in the electricity markets compared to traditional financial markets, one often resorts to the most liquid derivatives when calibrating models, meaning forward contracts[1] in the Europan electricity markets. The unstorable nature of electricity implies incompleteness of electricity markets. As in e.g. Benth & Meyer-Brandis (2009), we therefore define the forward price at time s with delivery time t as $F(s,t) = \mathbb{E}^{\mathbb{Q}}[S(t)|\mathcal{F}_s]$. Using parameterized models as the ones from Margrabe (1978) and Christensen & Benth (2019), the recipe we follow in order to price a transmission right is therefore to force the models to be consistent with the forward prices. That is, each hourly model (24 models in total for pricing and hedging a transmission right) will be calibrated to the observed forward prices. Afterwards, the valuation of the individual spread options, $(g^\top S(t,h))^+$, are added up cf. Eq. (E.1), giving the resulting estimated transmission right price.

2.1 Hedging Transmission Rights

We limit the available hedging instruments to baseload forward contracts. In reality, liquidity issues also plays a role in picking the hedging portfolio, and since the baseload forward contracts are quite liquid compared to other instruments in the electricity market, this is a reasonable assumption. Thus, hedging e.g. a monthly transmission right means taking positions in the underlying available forward contracts, which could be with a delivery period of one day, a week, etc. Taking into account this overlapping nature of the forward contracts and the tick size of the contracts are important when

[1] We assume a deterministic interest rate, hence, we do not distinguish between futures or forward contracts. We only use the term forward contracts in the rest of the paper.

hedging transmission rights. Mathematically, it can be described by

$$\min_{Q(1),..,Q(M)} \sum_{t=1}^{T} \sum_{h=1}^{H(t)} \sum_{i=1}^{M} \mathbb{I}(i,t,h) \left(X(t,h) + Q(i)\right)^2, \quad Q(i) \in \mathcal{K} \quad \forall i = \{1,...,M\}. \tag{E.2}$$

Here, $\mathbb{I}(i,t,h)$ is an indicator function being one if the delivery period of the ith instrument overlaps in time with day t and hour h, $X(t,h)$ is the exposure from the transmission right on day t and hour h, and M denotes the number of forward contracts used to hedge. $\mathcal{K}$ is the set of allowed quantities to trade; e.g., contracts can only be traded in an integer amount of MWs. Lastly, $Q(i)$ is the traded quantity in MW of the ith instrument.

3 Model Description

In this section, we present the Margrabe formula [Margrabe (1978)] followed by a presentation of the Conv model from Christensen & Benth (2019) in a transmission right pricing and hedging context.

3.1 The Margrabe Formula

Denote the electricity day-ahead price for two interconnected markets by $S(t) = [S_1(t) \quad S_2(t)]^\top$. The Margrabe formula is based on the assumption that the underlyings follow a geometric Brownian motion under a pricing measure Q,

$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t) dW_i(t), \quad i = 1,2,$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2 \in \mathbb{R}$, and $W_1(t)$ and $W_2(t)$ are correlated Brownian motions with correlation $\text{corr}(W_1(t), W_2(t)) = \rho$. The Margrabe formula for a spread option with payoff $\Pi(s,t) := \mathbb{E}^Q[(S_1(t) - S_2(t))^+ | \mathcal{F}_s]$ is given by

$$\Pi(s,t) = S_1(s)e^{\mu_1(t-s)}\Phi(d_1) - S_2(s)e^{\mu_2(t-s)}\Phi(d_2), \tag{E.3}$$

where

$$d_1 = \frac{\ln\left(\frac{S_1(s)}{S_2(s)}\right) + \left(\mu_1 - \mu_2 + \frac{\sigma^2}{2}\right)(t-s)}{\sigma\sqrt{t-s}},$$

$$d_2 = d_1 - \sigma\sqrt{t-s},$$

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$

It is well-known that the geometric Brownian motion implies that the forward price is given by

$$F_i(s,t) = S_i(s)e^{\mu_i(t-s)}, \quad i = 1,2.$$

Hence, in order to calibrate the Margrabe formula to the forward prices, one simply chooses μ_i so the model forward price matches the observed forward price. It can easily be shown that this valuation approach is equivalent to using the observed forward prices in a variation of Eq. (E.3); that is,

$$\Pi(s,t) = F_1(s,t)\Phi(d_1^F) - F_2(s,t)\Phi(d_2^F), \tag{E.4}$$

where

$$d_1^F = \frac{\ln\left(\frac{F_1(s,t)}{F_2(s,t)}\right) + \frac{\sigma^2}{2}(t-s)}{\sigma\sqrt{t-s}},$$

$$d_2^F = d_1^F - \sigma\sqrt{t-s}.$$

In this way, the model will be consistent with the observed forward prices.

Hedging

The approach in Eq. (E.4) has an advantage over Eq. (E.3) in a hedging context: Since we use forward contracts to hedge transmission rights, sensitivities can directly be calculated with regard to the forward prices instead of the day-ahead prices. Let $\Pi(s,t)$ denote the transmission right price at time s for delivery at time t. The derivatives of $\Pi(s,t)$ with respect to both $F_1(s,t)$ and $F_2(s,t)$, i.e. the deltas of the transmission right, are

$$\begin{aligned}\Delta_1 &:= \frac{\delta\Pi(s,t)}{\delta F_1(s,t)} = \Phi(d_1^F),\\ \Delta_2 &:= \frac{\delta\Pi(s,t)}{\delta F_2(s,t)} = -\Phi(d_2^F).\end{aligned}$$

3.2 The Conv Model

The following is a brief introduction to the Conv model. The reader is referred to Christensen & Benth (2019) for more information and for the proofs of the presented propositions.

We assume that the day-ahead electricity prices at time t in $N = 2$ interconnected market areas can be described by

$$S(t) = I(t)p(t), \tag{E.5}$$

where $I(t)$ takes two forms,

$$I(t) = \begin{cases} I_{(1)}, & \text{if } z(t) > 0 \\ I_{(2)}, & \text{otherwise} \end{cases},$$

$$I_{(1)} = \begin{bmatrix} w & 1-w \\ w & 1-w \end{bmatrix} \text{ and } I_{(2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{E.6}$$

depending on a process $z(t)$ defined by

$$\begin{aligned} z(t) &= \lambda(t) + x(t), \\ dx(t) &= -kx(t)dt + \sigma(t)dB_z(t). \end{aligned}$$

$k \in \mathbb{R}$ is a constant, and $B_z(t)$ is a univariate Brownian motion, while $\lambda(t)$ and $\sigma(t)$ are deterministic functions. $I_{(1)}$ represents the EPC state, and $I_{(2)}$ represents the non-EPC state, where the prices in the two areas will simply be $p(t)$. Hence, the interpretation of $p(t)$ is that it represents the partly latent "domestic" day-ahead prices. In Eq. (E.6), $w \in (0,1)$ measures the pulling effect of each market area. In other words, if EPC is observed, which market area influences the EPC price the most.

Turning again to Eq. (E.5), $p(t)$ is defined as

$$p(t) = \Lambda(t) + q(t), \tag{E.7}$$

$$dq(t) = -Kq(t)dt + \Sigma dB_p(t). \tag{E.8}$$

In Eq. (E.7) and Eq. (E.8), $\Lambda(t)$ is a deterministic function, and $K \in \mathbb{R}^{2\times 2}$ is a constant matrix, $\Sigma \in \mathbb{R}^{2\times 2}$, and $B_p(t)$ is an N-dimensional Brownian motion. The conditional distribution of $S(t)$ under a pricing measure Q is given in Prop. 8. Notice the use of the stack operator denoted by $\text{vec}(\cdot)$, and the Kronecker sum between the two square matrices A and B defined by $A \oplus B = A \otimes \mathbf{I}_B + \mathbf{I}_A \otimes B$, where $\otimes$ is the Kronecker product. Here $\mathbf{I}_B$ is the identity matrix with the same size as B, and $\mathbf{I}_A$ is the identity matrix of same size as A.

Proposition 8
Let $p(t)$ and $z(t)$ be defined as above. Assume no market price of risk regarding the $z(t)$ process. The distribution of $S(t)$ is then mixed normal conditional on $\mathcal{F}_s$ for $s < t$ under a pricing measure Q with density

$$f^{Q}_{s,t}(x) = \sum_{j=1}^{2} \mathbb{P}_{(j)}(s,t) \frac{1}{(2\pi)^{N_j/2}|\Sigma_{(j)}(s,t)|} e^{-\frac{1}{2}\left(x-\mu^{Q}_{(j)}(s,t)\right)^{\mathrm{T}}\Sigma_{(j)}(s,t)^{-1}\left(x-\mu^{Q}_{(j)}(s,t)\right)},$$

with $N_j \leq N$ for $j = \{1,2\}$ being the rank of $I_{(j)}$, $\Phi(\cdot)$ being the cumulative distribution function for the standard normal distribution, and

$$\begin{aligned}\mu^Q_{(j)}(s,t) =& I_{(j)}\mathbb{E}^Q\left[p(t)|\mathcal{F}_s\right],\\ \Sigma_{(j)}(s,t) =& I_{(j)}\mathbb{V}\mathrm{ar}[p(t)|\mathcal{F}_s]I^{\mathrm{T}}_{(j)},\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}^Q\left[p(t)|\mathcal{F}_s\right] &= \Lambda(t) + e^{-(t-s)K}q(s) + \int_s^t e^{-(t-u)K}\Sigma\varphi(u)du,\\ \mathbb{V}\mathrm{ar}[p(t)|\mathcal{F}_s] &= \mathrm{vec}^{-1}\left((K\oplus K)^{-1}(\mathbf{I} - e^{-(t-s)K\oplus K})\mathrm{vec}(\Sigma\Sigma^{\mathrm{T}})\right),\end{aligned}$$

where $\varphi(t) \in \mathbb{R}^2$ is the market price of risk. Furthermore, the probabilities

$$\mathbb{P}_{(1)}(s,t) = 1 - \Phi\left(\frac{-\mathbb{E}[z(t)|\mathcal{F}_s]}{\sqrt{\mathbb{V}\mathrm{ar}[z(t)|\mathcal{F}_s]}}\right), \quad \mathbb{P}_{(2)}(s,t) = 1 - \mathbb{P}_{(1)}(s,t).$$

As shown in Christensen & Benth (2019), Prop. 8 implies the following two propositions.

Proposition 9
Let $S(t)$ be defined under Q as in in Prop. 8. Then the forward price at time s for delivery at time t for two different market areas is

$$F(s,t) = \left[\mathbb{P}_{(1)}(s,t)I_{(1)} + \mathbb{P}_{(2)}(s,t)I_{(2)}\right]\mathbb{E}^Q\left[p(t)|\mathcal{F}_s\right]. \tag{E.9}$$

Proposition 10
Let $S(t)$ be defined under Q as in Prop. 8, $g \in \mathbb{R}^2$, and let $(g^{\mathrm{T}}S(t))^+$ be the payoff of a European spread option. Then

$$\begin{aligned}\Pi(s,t) &= \mathbb{E}^Q[(g^{\mathrm{T}}S(t))^+|\mathcal{F}_s]\\ &= \sum_{i=1}^{2}\mathbb{P}_{(i)}(s,t)\left(a_{(i)}(s,t)\Phi\left(\frac{a_{(i)}(s,t)}{c_{(i)}(s,t)}\right) + c_{(i)}(s,t)\phi\left(\frac{a_{(i)}(s,t)}{c_{(i)}(s,t)}\right)\right),\end{aligned} \tag{E.10}$$

where

$$a_i(s,t) = g^{\mathrm{T}}\mu^Q_{(i)}(s,t),$$

$$c_i(s,t) = \sqrt{g^{\mathrm{T}}\Sigma_{(i)}(s,t)g}, \tag{E.11}$$

and Φ and ϕ is the distribution and density function of the standard normal distribution, respectively. $\mu^Q_{(i)}(s,t)$ and $\Sigma_{(i)}(s,t)$ is given in Prop. 8.

In Eq. (E.10) the forward price is not present, but rather the "domestic" forward price $\mathbb{E}^{\mathbb{Q}}[p(t)|\mathcal{F}_s]$. However, if $\mathbb{P}_{(2)}(s,t) > 0$, which is a reasonable assumption having the discussion on the market coupling in Sec. 1 in mind, we get from Eq. (E.9) that

$$\mathbb{E}^{\mathbb{Q}}[p(t)|\mathcal{F}_s] = \left[\mathbb{P}_{(1)}(s,t)I_{(1)} + \mathbb{P}_{(2)}(s,t)I_{(2)}\right]^{-1} F(s,t).$$

Hence, the observed forward prices in the market can directly be converted to the domestic forward prices, needed in Eq. (E.10). This is useful in both a pricing and hedging setting. Going forward, we refer to $\mathbb{E}^{\mathbb{Q}}[p(t)|\mathcal{F}_s]$ as the domestic forward price.

One can verify that the spread between the forward prices, $F_1(s,t) - F_2(s,t)$ where $F(s,t) = [F_1(s,t) \quad F_2(s,t)]^\top$, does not depend on w in $I_{(1)}$ but only on the probability of EPC, meaning that the spread between the domestic forward prices does not depend on w as well. In fact,

$$\mathbb{E}^{\mathbb{Q}}[p_1(t) - p_2(t)|\mathcal{F}_s] = \frac{F_1(s,t) - F_2(s,t)}{\mathbb{P}_{(2)}(s,t)}, \tag{E.12}$$

where $p(t) = [p_1(t) \quad p_2(t)]^\top$. For $\mathbb{P}_{(2)}(s,t) = 1$, the domestic forward price spread will coincide with the observed forward price spread as expected, while $\mathbb{E}^{\mathbb{Q}}[p_1(t) - p_2(t)|\mathcal{F}_s] > [F_1(s,t) - F_2(s,t)]$ for $\mathbb{P}_{(2)}(s,t) \in (0,1)$. As $\mathbb{P}_{(2)}(s,t) \to 0$ we have that $\mathbb{E}^{\mathbb{Q}}[p_1(t) - p_2(t)|\mathcal{F}_s] \to \infty$. In other words, the smaller $\mathbb{P}_{(2)}(s,t)$ is, the more sensitive the domestic forward price spread is towards changes in the observed forward price spread. We return to the implication of this feature on transmission rights shortly.

Lastly we note that one typically has that $g = [1 \quad -1]^\top$ in a transmission right context (which is the case considered in the remaining part of this paper), reducing Eq. (E.10) to

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[(S_1(t) - S_2(t))^+|\mathcal{F}_s] =& \mathbb{P}_{(2)}(s,t)\left(a_{(2)}(s,t)\Phi\left(\frac{a_{(2)}(s,t)}{c_{(2)}(s,t)}\right)\right. \\ &\left. + c_{(2)}(s,t)\phi\left(\frac{a_{(2)}(s,t)}{c_{(2)}(s,t)}\right)\right).\end{aligned} \tag{E.13}$$

The transmission right price in Eq. (E.13) is also independent of the choice of w.

Hedging

In Prop. 11 we present results regarding the delta sensitivities implied by the Conv model.

Proposition 11
Let $S(t)$ be defined under Q as in Prop. 8, $g \in \mathbb{R}^2$, and let $(g^{\mathrm{T}}S(t))^+$ be the payoff of a European spread option. Define

$$J_{(i)}(s,t) = I_{(i)} \left[\mathbb{P}_{(1)}(s,t) I_{(1)} + \mathbb{P}_{(2)}(s,t) I_{(2)} \right]^{-1}.$$

Then

$$\begin{aligned} \begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix} &= \frac{\partial \Pi(s,t)}{\partial F(s,t)} \\ &= \sum_{i=1}^{2} \mathbb{P}_{(i)}(s,t) \Phi \left(\frac{a_{(i)}(s,t)}{c_{(i)}(s,t)} \right) J_{(i)}(s,t)^\top g, \end{aligned}$$

where $a_i(s,t)$ and $c_i(s,t)$ are defined in Prop. 8.

Proof. Notice that

$$a_i(s,t) = g^\top J_{(i)}(s,t) F(s,t).$$

Hence,

$$\frac{\partial a_i(s,t)}{\partial F(s,t)} = J_{(i)}(s,t)^\top g,$$

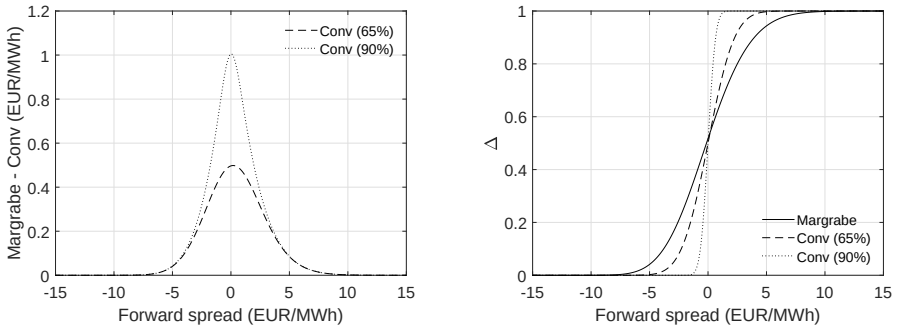
and the rest follows by straightforward calculations. □

Assuming $g = [1 \quad -1]^\top$, the delta is simplified to

$$\frac{\partial \Pi(s,t)}{\partial F(s,t)} = \Phi \left(\frac{a_{(2)}(s,t)}{c_{(2)}(s,t)} \right) g. \tag{E.14}$$

3.3 Comparison of Margrabe and Conv

In Fig. E.4(a) we consider the difference in the transmission right price for a single hour as a function of different forward spreads. The parameters used in Fig. E.4(a) are all taken from the estimation results reported in Sec. 4.1. The estimated parameters for the $z(t)$ process translate to a probability of EPC of roughly 65% in the Conv model. The prices implied by the two models differ mostly at-the-money (ATM), with the Margrabe formula pricing highest. The difference with regard to deep in-the-money (ITM) and deep out-of-the-money (OTM) is opposite quite small. The interpretation is that when the forward spread is large in absolute terms, the probability of EPC becomes less important; the forward spread simply dictates the transmission right price. By forcing an increase in the probability of EPC from 65% to 90%, the



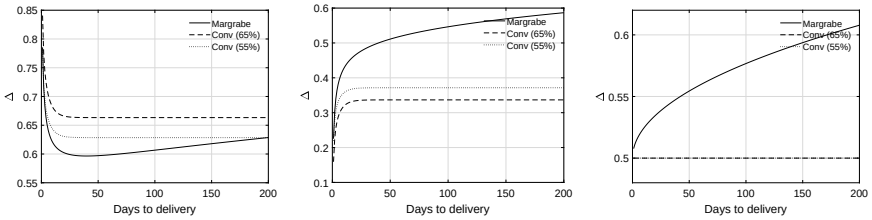
(a) Differences in transmission right prices

(b) Delta for the Margrabe and Conv model

Fig. E.4: a) difference between transmission right prices for a single hour implied by the Conv model and the Margrabe formula for a varying forward spread, b) the deltas of the long leg implied by the Margrabe formula and Conv model. In the Conv case, we use a probability of EPC of both 65% and 90%. The forward spread is calculated by $F_1 - F_2$, where $F_2 = 35$ EUR/MWh and F_1 varies between 20 EUR/MWh and 50 EUR/MWh. Time to delivery is 5 days.

difference of ATM transmission rights increases between the Margrabe and Conv model.

The delta sensitivities are depicted in Fig. E.4(b), showing the deltas, denoted by Δ, of the long leg of a transmission right using the same model specification as in Fig. E.4(a). To further illustrate the impact of the probability of EPC, we again show two versions of the Conv model: A model with a EPC probability of 65% and a model with a EPC probability forced to 90%. Comparing the Margrabe implied delta to the Conv implied delta reveals that the Conv model is more sensitive towards changes in the forward spread. By increasing the probability of EPC [decreasing $\mathbb{P}_{(2)}(s,t)$], we increase $a_{(2)}(s,t)$ in absolute terms in Eq. (E.14) as a consequence of Eq. (E.12), thereby leading to a more extreme value of Δ. Hence, the 90% Conv version is more extreme compared to the 65% version in the sense that it is more sensitive towards changes in the underlying. As the probability of EPC gets close to one, the Conv model suggests that the transmission right price approximates the behaviour of the positive forward spread; the delta is one if the spread is positive, and otherwise zero. To draw a parallel to the Margrabe formula, increasing the EPC probability in the Conv model has the same qualitative effect on the transmission right price as decreasing the volatility, σ, in the Margrabe formula. Obviously, the volatility and correlation of the domestic day-ahead prices influencing $c_{(2)}(s,t)$ also affects the level of the extrinsic value of the transmission right price implied by the Conv model. It can be read directly off from Eq. (E.14) that decreasing $c_{(2)}(s,t)$ will increase the sensitivity of the transmission right price towards changes in the forward spread.



(a) ITM (b) OTM (c) ATM

Fig. E.5: Δ as a function of time to delivery for a) ITM transmission rights b) OTM transmission rights, and c) ATM transmission rights. The employed models are the Margrabe formula and Covn model using the parameters reported in Sec. 4.1. The EPC probability in the Conv model is approximately 65%. A Conv model with a EPC probability forced to 55% is also shown.

Fig. E.5 shows Δ of the long leg as a function of time to delivery for ITM, OTM, and ATM transmission rights for both the Margrabe and Conv model. For illustrative purposes Fig. E.5 shows a Conv model with a EPC probability of 65% and another with a EPC probability forced to 55%. If the EPC probability decreases, then the spread between the domestic forward prices will also decrease as already discussed. Hence the decrease in Δ for ITM transmission right, and the increase in Δ for OTM transmission rights. For ATM transmission rights, the Δ is constant at 0.5 in both cases. Comparing the Conv model to the Margrabe formula, they both imply an increase/decrease in Δ for ITM/OTM transmission rights when the time to delivery is decreased. Due to the assumption of mean reversion in the Conv model of the domestic day-ahead prices, the Δ will not increase with time to delivery like the Margrabe formula implies.

4 Empirical Results

We consider in this section pricing and hedging of transmission rights using the two models presented in Sec. 3. From the Joint Allocation Office[2], the results of auctions for monthly transmission rights in Europe are available as well as the auction periods. In order not to disturb the results due to changing market environments regarding the way the market coupling is effectuated, such as the change from the available transfer capacity method to the flow-based one as discussed in Sec. 1, our observation period is from 21 July 2015 to 25 April 2019. However, we require two years of day-ahead price observations when estimating the models, so the delivery of the first transmission right we consider starts 1 August 2017, and the delivery period of the last transmission right starts 1 May 2019. We acknowledge that other

[2]See jao.eu.

changes in the market environment may also have an impact on the market coupling. As an example, the splitting of the German/Austrian day-ahead auction in October 2018 could potentially also impact the market coupling in the CWE area, even though the shift from the available transfer capacity method to the flow-based method has a more direct effect on the market coupling. A further analysis is outside the scope of present paper and is left as further research.

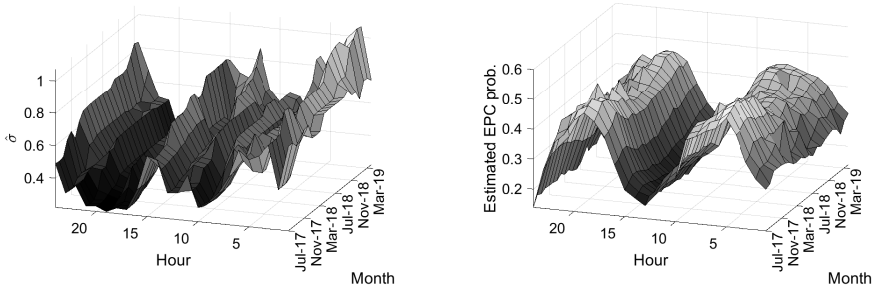
The considered period translates to 22 transmission rights for each cross-border connection. We consider the transmission rights DE→FR, FR→DE, DE→NL, and NL→DE, resulting in 88 observed transmission right auctions in total. DE→FR should be interpreted as the transmission right with a long French leg and a short German leg; i.e., Germany to France. The day-ahead prices can be obtained from the EPEX SPOT exchange, while the baseload forward prices used in the hedging application can be obtained from the European Energy Exchange.

In the following, we first consider estimation results. Secondly, we discuss the empirical pricing results of both the Margrabe formula and Conv model, and also compare these to the traded transmission right prices. This is followed by a discussion of the transmission right hedging ability of the two models in a delta hedging context.

4.1 Estimation

As discussed in Sec. 1, we exploit the hourly separation of the payoff structure of a transmission right displayed in Eq. (E.1) by modeling each hourly day-ahead price independently of all other hourly day-ahead prices. Each hourly model will be referred to as a submodel and it is estimated on historical day-ahead prices for the hour in question; e.g., the submodel for hour 1 (the hour from 00:00 to 01:00) will be estimated on historical hour 1 day-ahead prices.

In both the Margrabe and Conv case, we first employ historical day-ahead prices to estimate the needed parameters, and then forward prices to shift measure, as discussed in Sec. 3. We estimate the parameters for all 24 submodels in both cases on the last day of each auction period. This results in $2112 = 24 \times 22 \times 2 \times 2$ (number of hours, number of transmission right auctions, number of cross-border connections, number of models) estimated submodels in total. To keep the presentation as clear as possible, we therefore only show the estimated parameters on two years of hour 1 day-ahead price observations ending 21 July 2017 for the joint modeling of Germany and France. It corresponds to the last day of the auction period for the monthly transmission right with delivery period in August 2017. We complement these results by figures intended to give the reader an insight in the remaining estimation results.



(a) Margrabe (b) Conv

Fig. E.6: a) estimated $\hat{\sigma}$ for the Margrabe formula, and b) the mean of the asymptotic EPC probabilities throughout a year for the Conv model. In both cases results are shown for all hours and all estimation periods.

Margrabe Formula

The simplicity of the Margrabe formula imposes a quick and straightforward estimation procedure. As is well-known, see e.g. Campbell *et al.* (1996), one simply estimates σ_i by taking the standard deviation of the log returns of the day-ahead prices from the ith market area for the hour in question. ρ is estimated by the linear correlation coefficient between the same log return time series from the two market areas. The estimated parameters are $\hat{\sigma}_1 = 0.74$, $\hat{\sigma}_2 = 0.21$, and $\hat{\rho} = 0.15$. Fig. E.6(a) shows $\hat{\sigma} := \sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2 - 2\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2}$ for all hours and all estimation periods. A general pattern through the hour dimension can be observed with the minimum values of $\hat{\sigma}$ attained around hour 10 and 19, and peaks around hour 2, 15, and 24. There is an increasing tendency along the time dimension. Since transmission right prices implied by the Margrabe formula is increasing in σ, the Margabe formula thus implies increasing transmission right prices. By assuming that increased price convergence is equivalent to a decrease in the transmission right price, Fig. E.6(a) is consistent with the discussion in Sec. 1 and therefore violates the target of the European Commission of increased price convergence.

Conv Model

We refer the interested reader to Christensen & Benth (2019) for the details of the estimation procedure for the Conv model. To estimate w in Eq. (E.6), we use the supplied electricity of each market area. The used supplied electricity is given in Table E.2.

The estimation procedure for the Conv model is split up in a Kalman filtering and particle filtering part. Since the observed forward prices in the

DE	FR	NL
575,701	496,932	114,201

Table E.2: Supplied electricity given in (GWh) based on IEA (2017).

market areas determine $\mathbb{E}^{\mathbb{Q}}[p(t)|\mathcal{F}_s]$ cf. the discussion in Sec. 3, we will not dwell with the estimation results of $\Lambda(t)$, but only note that we use the same specification as in Christensen & Benth (2019). Moving on to the parameters affecting $c_i(s,t)$ in Eq. (E.11), we assume that Σ is lower triangular, and that K is diagonal in Eq. (E.8). In Table E.3 the estimation results for $q(t)$ are reported.

$\hat{K}_{(1,1)}$	$\hat{K}_{(2,2)}$	$\hat{\Sigma}_{(1,1)}$	$\hat{\Sigma}_{(2,1)}$	$\hat{\Sigma}_{(2,2)}$
85.72	35.54	82.72	54.25	52.41

Table E.3: Estimation results for $q(t)$.

Lastly we report the estimated parameters for the $z(t)$ process in Table E.4. Based on the discussion in Christensen & Benth (2019), we use the following specifications of $\lambda(t)$ and $\sigma(t)$:

$$\lambda(t) = a_\lambda + b_\lambda \cos\left(\frac{2\pi t}{M} + c_\lambda\right),$$
$$\sigma(t) = a_\sigma + b_\sigma \cos\left(\frac{2\pi t}{M} + c_\sigma\right),$$
$$a_\sigma = \sqrt{\frac{2k}{1 - e^{-2k(t-s)}}},$$

where $t - s$ measures the time between the equidistant observations. That is, we allow yearly seasonality in the probability of EPC and in the uncertainty of it. Here a_λ, b_λ, c_λ, k, a_σ, b_σ, $c_\sigma \in \mathbb{R}$ and $M = 365$. To give a sense

$\hat{k}$	$\hat{a}_\lambda$	$\hat{b}_\lambda$	$\hat{c}_\lambda$	$\hat{b}_\sigma$	$\hat{c}_\sigma$
131.51	-1.00	1.18	0.18	-19.47	-0.60

Table E.4: Estimation results for $z(t)$.

of the estimation results for all submodels, we calculate the asymptotic EPC probability for each day during a year implied by the corresponding submodel. Since we employ a different submodel for each hour, this procedure results in 365 asymptotic EPC probabilities for each hour. Fig. E.6(b) shows

the development of the mean of these asymptotic EPC probabilities for each hour through time. A general observation is that the EPC probabilities for all hours are fairly constant over time. The peaks in EPC probabilities are centred around hour 5 and hour 19, with minimums at hour 13 and hour 24. This pattern is observed through all estimation periods. Comparing the shape of the hour dimension of Fig. E.6(b) to Fig. E.6(a), minimums in Fig. E.6(b) are roughly attained at maximums in Fig. E.6(a). The EPC aspect of the Conv model is handled by the Margrabe formula by adjusting σ: A high probability of EPC corresponds to a low value of σ, and vice versa.

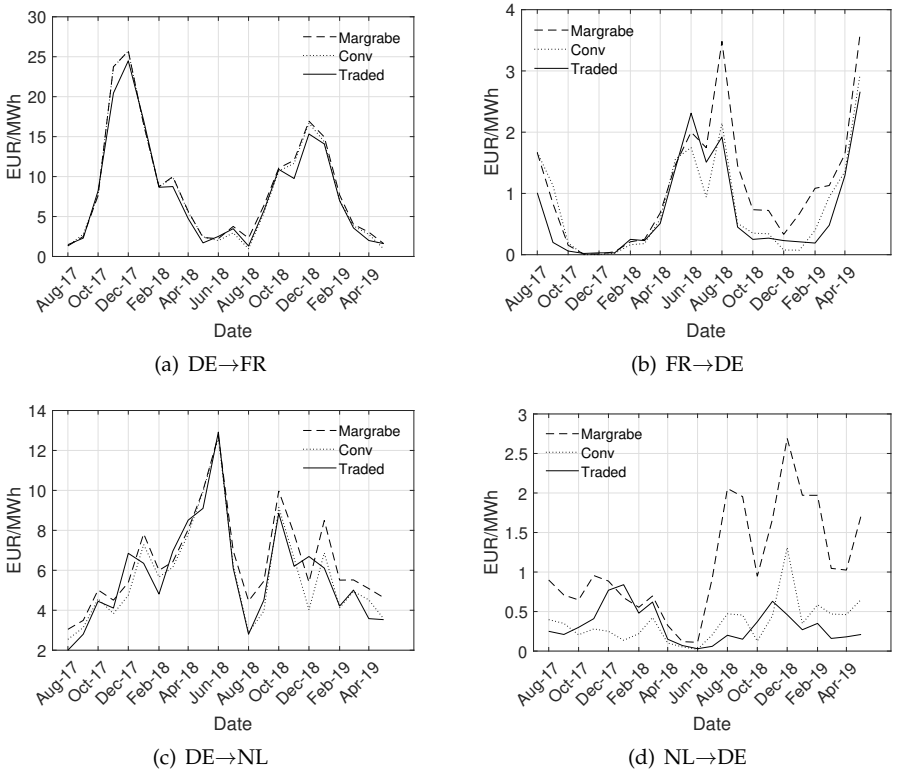
We finally note that the oscillation of the EPC probability between roughly 20-50% is in line with the number reported for the DE-FR cross-border in Table E.1 of 35.69%.

4.2 Pricing

Fig. E.7 shows the calculated transmission right prices on the last auction date using the Margrabe formula and the Conv model. For comparison, we also show the auctioned transmission right prices for each cross-border connection. Firstly, the two approaches seem to replicate the traded prices quite well, especially when valuing ITM transmission rights. On average the monthly forward spread in Fig. E.7(a) between France and Germany (that is, the German forward price subtracted from the French forward price) is 7.35 EUR/MWh on the auction days for the periods shown in Fig. E.7, while the average spread between the Netherlands and Germany in Fig. E.7(c) is 5.40 EUR/MWh. Notice that these are the settlement prices provided by EEX on auction dates. However, the OTM transmission rights (on average) illustrated in Fig. E.7(b) and Fig. E.7(d) indicates that the Margrabe formula struggles to replicate the market behaviour, whereas the Conv model is more in line with the market.

Fig. E.8 shows a two-dimensional histogram of the forward spread on auction dates and the corresponding differences between the Margrabe and Conv transmission right prices for all price spreads; i.e., the number of instances for a particular forward price spread and the corresponding spread between the Margrabe implied and the Conv implied transmission right price. The difference between the Margrabe formula and Conv model fades out the greater the forward spread is in absolute terms, confirming the conclusion from Fig. E.4(a). Notice in particular the similarity of the shapes of Fig. E.4(a) and Fig. E.8. In most cases, the Margrabe formula estimates the transmission right price to be higher than the corresponding Conv price; the price difference is on average 0.53 EUR/MWh. To set it in perspective, the average traded transmission right price is 3.68 EUR/MWh. The mean in the Margrabe and Conv case is 4.32 EUR/MWh and 3.80 EUR/MWh, respectively.

To further elaborate on the pricing differences between the Margrabe and



(a) DE→FR (b) FR→DE

(c) DE→NL (d) NL→DE

Fig. E.7: Calculated transmission rights prices on the last auction date and the corresponding auctioned prices, (a) transmission right from DE to FR, (b) transmission right from FR to DE, (c) transmission right from DE to NL, (d) transmission right from NL to DE.

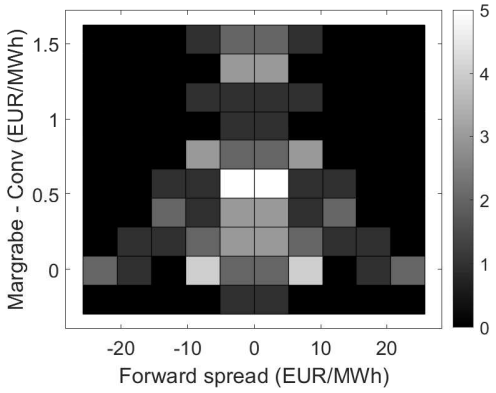


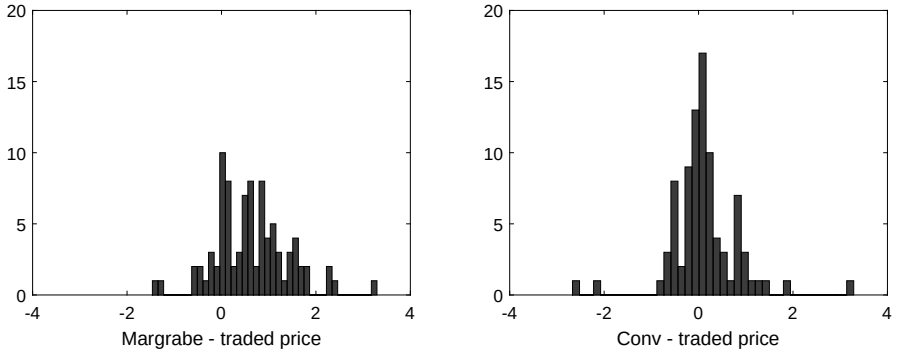
Fig. E.8: A two-dimensional histogram with the forward spreads on auction dates versus the differences between the estimated transmission right prices by the Margrabe formula and the Conv model.

Conv approach, the minimum, mean, and maximum of the difference between each model implied transmission right price and the corresponding traded price for each price spread are given in Table E.5. In all cases does the implied Margrabe price surpass the traded price on average, whereas the Conv model is more consistent with the market according to the mean values. The dispersion of the distributions measured by the average spread between the maximum and minimum values is 3.00 and 2.65 in the Margrabe and Conv case, respectively. Fig. E.9 shows a histogram of the differences

	DE→FR	**FR→DE**	**DE→NL**	**NL→DE**
		Margrabe		
Minimum	-0.31	-0.58	-1.46	-0.16
Mean	0.38	0.69	0.70	0.79
Maximum	1.56	3.29	2.39	2.23
		Conv		
Minimum	-0.57	-0.66	-2.67	-0.70
Mean	0.08	0.39	-0.056	0.04
Maximum	0.93	3.29	0.94	0.85

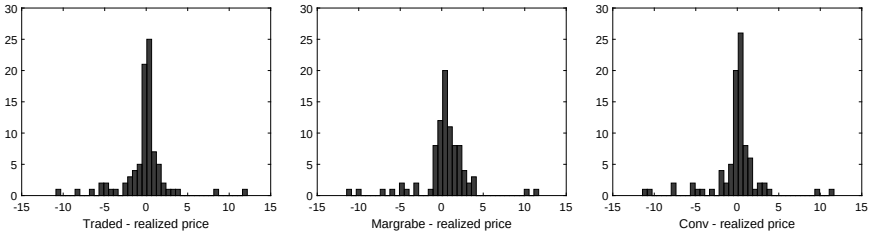
Table E.5: Minimum, mean, and maximum of the difference between the implied transmission right price and the corresponding traded price at the four cross-borders. The numbers are denoted in EUR/MWh and given for the two models, Margrabe and Conv.

between the prices implied by each model and the traded prices for all price spreads. The differences in Fig. E.9(a) is more spread out than in Fig. E.9(b), which is more centred around zero. Also, the mean of the differences in Fig. E.9(a) is 0.64 EUR/MWh compared to 0.11 EUR/MWh in Fig. E.9(b).



(a) Margrabe (b) Conv

Fig. E.9: a) the difference between the transmission right price implied by the Margrabe formula and the corresponding traded price, and b) the difference between the transmission right price implied by the Conv model and the corresponding traded price.



(a) Traded (b) Margrabe (c) Conv

Fig. E.10: a) the difference between the traded and realized transmission right prices, and b) the difference between the transmission right prices implied by the Margrabe formula and the realized prices, and c) the difference between the transmission right prices implied by the Conv model and the realized prices.

It is also interesting to consider the prices at which the transmission rights were realized; that is, the payoff from holding the transmission right through the corresponding delivery period. Fig. E.10 compares the realized transmission right prices to the traded, Margrabe, and Conv prices. On average, the transmission right prices are traded at a lower price compared to the realized values, with a mean of the difference between the traded and realized prices being -0.23 EUR/MWh. Consequently, the difference between the Margrabe and realized prices is also smaller than the difference between the Margrabe and traded prices. The mean in Fig. E.10(b) is 0.41 EUR/MWh. The Conv model is the one closest to the realized prices, with -0.12 EUR/MWh being the average in Fig. E.10(c). Also notable is the more centred distribution around zero in Fig. E.10(a) and Fig. E.10(c) compared to Fig. E.10(b).

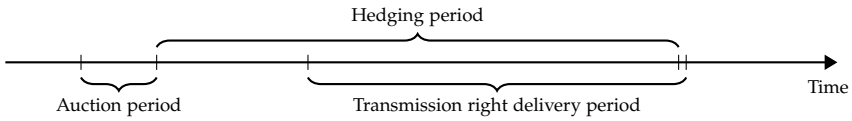
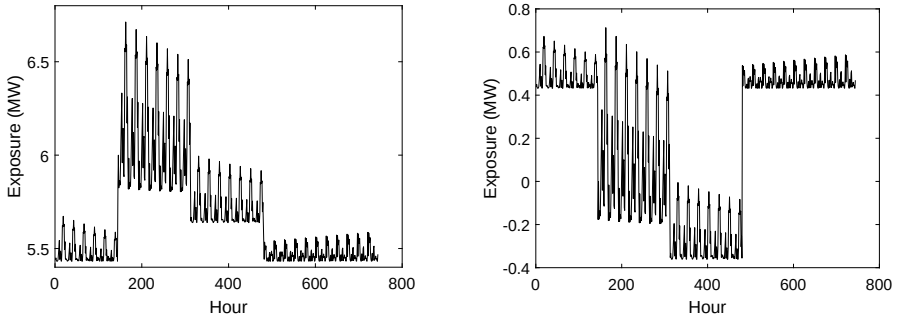


Fig. E.11: Time line overview of the hedging application.

4.3 Delta Hedging

In this section we consider the hedging aspect of transmission rights trading by employing delta hedging strategies imposed by the Margrabe formula and Conv model. We use the estimated parameters on end auction dates throughout each hedging period. Each hedging period is defined as the period from the last auction date to the second-to-last date in the delivery period of the transmission right (we cannot hedge on the last day) as schematically illustrated in Fig. E.11. For each transmission right, we re-evaluate our hedge daily during the hedging period. One might notice the spread in time between the auction period and the delivery period of the transmission right in Fig. E.11, where we also re-adjust our hedge according to the changes in the forward prices. Forward contracts are only traded on EEX on business days, and we therefore consider between 20-30 business days in each hedging period. In all cases, we use the settlement price for each business day to represent the price for trading a particular forward contract on exactly that day. We furthermore assume that the hedger has access to the daily, weekly, and monthly baseload forward contracts on each business day quoted by EEX, that the tick size is 1 MW, that the traded transmission right size is 10 MW, and zero transaction costs.

The hourly delta exposure from the transmission right is calculated using the before mentioned baseload forward contracts. In other words, a single forward price is used for all hourly options making up the transmission right, cf. Eq. (E.1). In case of overlapping contracts, e.g. a weekly and a monthly forward contract, we use the price from the contract with the shortest delivery period due to the assumption of a more representative price for the shorter contract. To exemplify, we show in Fig. E.12(a) the delta exposure for the German leg of the transmission right FR$\rightarrow$DE implied by the Margrabe formula. The exposure is calculated on the last day of the auction period, which is 21 July 2017. On this date, EEX quoted the monthly contract with delivery period in August, the weekly contract with delivery period starting 7 August 2017, and lastly the weekly contract with delivery period starting 14 August 2017. As a result, the delta exposure covering these two weeks is calculated differently with regard to the used forward prices. Hence, the shift in the level of the exposure. Since a different model is applied for each hour of the day, the exposure will also differ. By considering Eq. (E.2), we find



(a) Margrabe exposure before delta hedging (b) Margrabe exposure after delta hedging

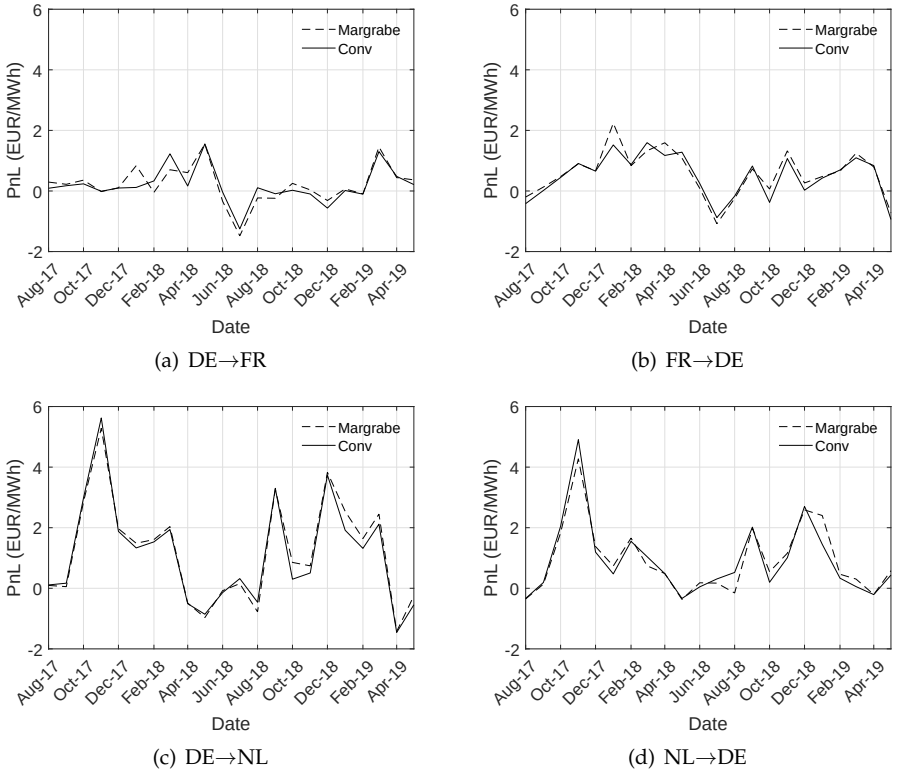
Fig. E.12: Delta exposure of the transmission right (a) before, and (b) after delta hedging for each hour of the delivery period of the transmission right delivering in August 2017. The exposure is calculated on the last day of the auction period, 21 July 2017. The underlying model is Margrabe.

that the optimal hedging quantities of the three forward contracts are -5 MW for the monthly contract, and -1 MW for both of the weekly contracts. The exposure for each hour of the transmission right delivery period after delta hedging with these quantities is shown in Fig. E.12(b). That is, Fig. E.12(b) shows the hourly exposure of the portfolio consisting of the 10 MW transmission right, the -5 MW monthly forward contract, and the two -1 MW weekly forward contracts. It is apparent that it is not possible to close down the exposure completely in this case, partly due to the constrained tick size and partly due to the fact that different submodels are applied to different hours. The difference in the submodels basically boils down to different values of σ in Eq. (E.4). Also, since the transmission right is ITM throughout the transmission right delivery period, the shape of each individual hour follows the shape of the Margrabe case in Fig. E.5(a).

The hedging results are summarized in Fig. E.13 and Table E.6. The profit

	DE→FR	**FR→DE**	**DE→NL**	**NL→DE**
		Margrabe		
Minimum	-1.48	-1.08	-1.42	-0.37
Mean	0.21	0.58	1.22	0.93
Maximum	1.55	2.24	5.28	4.28
		Conv		
Minimum	-1.25	-0.94	-1.46	-0.34
Mean	0.18	0.50	1.14	0.91
Maximum	1.55	1.60	5.62	4.91

Table E.6: Minimum, mean, and maximum of the PnL distribution implied by the delta hedging strategy of the transmission rights at the four cross-borders. The numbers are denoted in EUR/MWh and given for the two models, Margrabe and Conv.



(a) DE→FR (b) FR→DE

(c) DE→NL (d) NL→DE

Fig. E.13: Hedging results represented by the PnL given in EUR/MWh for (a) transmission rights from DE to FR, (b) transmission rights from FR to DE, (c) transmission rights from DE to NL, (d) transmission rights from NL to DE.

and loss (PnL) refers to the PnL generated by the portfolio of forward contracts used to hedge the transmission right and the cost of buying the corresponding transmission right. The delta hedging strategy implies a positive PnL on average in both the Margrabe and Conv case. The performance of both models is quite similar, which might be somewhat surprising. One explanation might be the tick size of 1 MW that implies that the differences between the hedges implied by the Margrabe formula and the Conv model are minimized. As an example, the Conv equivalent hedging portfolio to the Margrabe case shown in Fig. E.12 consists of -5 MW, -1 MW, and 0 MW for the monthly and two weekly forward contracts, respectively; i.e., in this case the Margrabe formula and the Conv model only differ by 1 MW on a weekly forward contract.

5 Conclusion

The pricing and hedging of transmission rights in coupled interconnected electricity markets is investigated in a continuous-time setting. Two models are employed, namely the Conv model from Christensen & Benth (2019) and the Margrabe formula from Margrabe (1978). The Conv model takes the market coupling into account and allows the exact price convergence observed empirically. Opposite, the Margrabe formula acts as a benchmark due to its well-known properties, but does not take the market coupling into account.

In a pricing context, we consider several monthly auctioned transmission rights in the Central-West Europe area. The Conv model seems to be more consistent with the market than the Margrabe formula. The former has a pricing error compared to the actual traded prices of 0.11 EUR/MWh on average, while the latter has a pricing error of 0.64 EUR/MWh. Also the distribution of the pricing errors of the Conv model is more evenly centred around zero compared to the Margrabe formula.

In a hedging context, a delta hedging strategy using both models is applied with baseload forward contracts of different delivery periods being the hedging instruments. The monthly transmission rights are each hedged from the last day of the auction period, where the position in the transmission right is also taken, to the second-to-last day of the delivery period of the transmission right. We find that the difference between hedging results implied by the two models is small.

Acknowledgements

The authors would like to thank Esben Høg and the Quantitative Analytics team at Centrica Energy Trading for valuable discussions. We also thank the participants attending the Vienna Congress on Mathematical Finance 2019 for fruitful discussions.

Disclosure Statement

No conflict of interest to declare.

Funding

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B.

References

Alasseur, C., Féron, O. (2018). Structural price model for coupled electricity markets. *Energy Economics*, **75**, 104–119.

Amprion. (2018). Flow based market coupling. Development of the market and grid situation 2015-2017.

Bachelier, L. (1900). Théorie de la spéculation. *Ann. Sci. Écol. Normale Supér.*, **17**, pp. 21–86.

Benth, F. E., Meyer-Brandis, T. (2009). The information premium for non-storable commodities. *Journal of Energy Markets*, **2**, 111–140.

Borovkova, S., Schmeck, M. D. (2017). Electricity price modeling with stochastic time change. *Energy Economics*, **63**, pp. 51–65.

Campbell., J. Y., Lo, A. W., MacKinlay, A. C. (1996). The Econometrics of Financial Markets. *Princeton University Press*, ISBN: 9781400830213.

Christensen, T. S., Benth, F. E. (2019). Modeling the joint behaviour of electricity prices in interconnected markets. *In review. See* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3312602` *for the working paper version*.

de Menezes, L. M., Houllier, M. A. (2016). Reassessing the integration of European electricity markets: A fractional cointegration analysis. *Energy Economics*, **53**, pp. 132–150.

den Bergh, K. V., Boury, J., Delarue, E. (2015). The Flow-Based Market Coupling in Central Western Europe: Concepts and definitions. *The Electricity Journal*, **29**, pp. 24–29.

Füss, R., Mahringer, S., Prokopczuk, M. (2015). Electricity market coupling in Europe: status quo and future challenges. *Working paper on finance no. 2015/12, University of St. Gallen*, **14**.

Geman, H., Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *Journal of Business*, **79**, pp. 1225–1261.

Huisman., R., Huurman, C., Mahieu, R. (2007). Hourly electricity prices in day-ahead markets. *Energy Economics*, **29**, pp. 240–248.

International Energy Agency (2017). Monthly electricity statistics. December 2017 report for electricity production and trade data for all OECD member countries.

Kiesel, R., Kustermann, M. (2015). Structural models for coupled electricity markets. Working paper. *SSRN*

Lucia, J. J., Schwartz, E. (2002). Electricity prices and power derivatives: evidence from the Nordic power exchange. *Review of Derivatives Research*, **5**, pp. 5–50.

Mahringer, S., Füss, R., Prokopczuk, M. (2015). Electricity market coupling and the pricing of transmission rights: an option-based approach. *Working papers on Finance 1512. University of St. Gallen, School of Finance*.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance*, **33**, pp. 177–186.

Pircalabu, A., Benth, F. E. (2017). A regime-switching copula approach to modeling day-ahead prices in coupled electricity markets. *Energy Economics*, **68**, pp. 283–302.

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-569-7 AALBORG UNIVERSITY PRESS